Aalto University
School of Chemical
Technology

School of Chemical Technology
Degree Programme of Chemical Technology

Jasmin Wang

CHARACTERIZATION AND ENGINEERING OF HEXARIC AND HEXURONIC ACID PATHWAYS IN FUNGAL MICROORGANISMS

Master's thesis for the degree of Master of Science in Technology submitted for inspection, Espoo, May 23rd 2016.

Supervisor Associate Professor Alexander Frey

Instructor Docent Peter Richard

Aalto University
School of Chemical
Technology

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of master's thesis

Author	Jasmin Wang
Title of thesis	Characterization and engineering of hexaric and hexuronic acid pathways in fungal microorganisms
Department	Biotechnology and Chemical technology
Professorship	Bioprocess engineering and food engineering
Code of professorship	KE-70
Thesis supervisor	Associate Professor Alexander Frey
Thesis advisor(s) / Thesis examiner(s)	Docent Peter Richard
Date	23.05.2016
Number of pages	70+11
Language	English

Abstract

Citrus juice production and sugar beet processing produce pectin-rich residues which are mainly used as animal feed. Food industry also uses pectin produced from citrus peel as gelling agent. The main constituent of pectin is D-galacturonic acid. The filamentous fungus *Aspergillus niger* is naturally capable of hydrolyzing pectin and it can catabolize galacturonic acid through pyruvate and glycerol. Hence, pectin-rich residues could provide a cost effective alternative for producing value-added products, for example, galactaric acid which has many potential applications in polymer synthesis and as a platform chemical. Galactaric acid could be produced by engineered *A. niger*. However, the acid is catabolized by the mould resulting in the lowered yield. This catabolic pathway of galactaric acid is still not known. The objectives of the Master's thesis were to unravel the pathway for galactaric acid catabolism, characterize the proteins and construct a galactaric acid producing *A. niger* strain.

In this study, eight potential genes involved in the catabolic pathways were investigated. The genes were deleted from the wild-type strain ATCC 1015 utilizing CRISPR/Cas9 technology, and galactaric acid utilization of the resulting mutant strains was investigated. To the best of my knowledge, this is the first report of using *in vitro* gRNA successfully in *A. niger* for targeted gene deletion by the CRISPR/Cas9 genome editing tool. Three of the investigated genes, *39114*, 1090836 and *1121140,* were concluded to be involved in the pathway. Blocking galacturonic and galactaric acid pathways by deleting genes *gaaA* and *39114*, respectively, and expressing a heterologous gene *udh*, *A. niger* Δ*gaaA udh* Δ*39114* was able to produce 4.3 ± 0.2 g/l galactaric acid from galacturonic acid in submerged fermentation. In solid-state and submerged fermentation, the strain produced 15.4 ± 0.6 and 82.3 ± 3.9 mg galactaric acid/g citrus processing waste on dry mass basis, respectively.

Keywords D-galacturonic acid, galactaric acid, citrus peel, filamentous fungi, *Aspergillus niger*, CRISPR/Cas9, homologous recombination

Aalto-yliopisto
Kemian tekniikan
korkeakoulu

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi

Diplomityön tiivistelmä

Tekijä Jasmin Wang		
Työn nimi Heksaari- ja heksuronihapporeittien karakterisointi ja muokkaus mikrosienissä		
Laitos Biotekniikan ja kemiantekniikan laitos		
Professuuri Biotekniikka ja elintarviketekniikka		**Professuurikoodi** KE-70
Työn valvoja Professori Alexander Frey		
Työn ohjaaja(t)/Työn tarkastaja(t) Dosentti Peter Richard		
Päivämäärä 23.05.2016	**Sivumäärä** 70+11	**Kieli** Englanti

Tiivistelmä

Sitrusmehun tuotannossa ja sokerijuurikkaan käsittelyssä muodostuu runsaasti pektiiniä sisältävää jätettä, jota pääsääntöisesti käytetään rehuna. Elintarviketeollisuus käyttää myös sitrushedelmien kuorista valmistettua pektiiniä hyytelöimisaineena. Pektiinin pääkomponentti on D-galakturonihappo. Filamenttihome *Aspergillus niger* kykenee luonnostaan hydrolysoimaan pekniitiä ja katabolisoimaan galakturonihapon pyruvaatiksi ja glyseroliksi. Näin ollen runsaasti pektiiniä sisältävät jätteet voisivat tarjota taloudellisesti kannattavan vaihtoehdon arvokkaiden tuotteiden, esimerkiksi galaktaarihapon, valmistukselle. Galaktaarihapolla on monia potentiaalisia sovelluksia polymeerisynteesissä. Geenimuokatulla *A. niger* –homeella voidaan tuottaa galaktaarihappoa, mutta home myös kykenee katabolisoimaan galaktaarihappoa, mikä alentaa saantoa. Galaktaarihapon kataboliareitti *A. niger* –homeessa ei kuitenkaan ole tunnettu. Tämän diplomityön tarkoituksena oli selvittää galaktaarihapon kataboliareitti, karakterisoida reittiin kuuluvat entsyymit ja konstruoida galaktaarihappoa tuottava *A. niger* –kanta.

Tässä työssä tutkittiin kahdeksan geeniä, jotka ovat mahdollisesti osana galaktaarihapon kataboliareittiä. Nämä geenit deletoitiin villityyppikannasta ATCC 1015 käyttämällä CRISPR/Cas9-teknologiaa. Parhaimman tiedon mukaan tämä on ensimmäinen raportti, jossa on käytetty *in vitro* gRNA:ta onnistuneesti geenien deletoimiseen CRISPR/Cas9 genomin muokkaustekniikalla. Kolme tutkituista geeneistä, *39114*, *1090836* ja *1121140*, todettiin olevan osana kataboliareittiä. Deletoimalla geenit *gaaA* ja *39114*, jotka ovat galakturoni- ja galaktaarihapon kataboliareiteissä, ja ekpressoimalla heterologinen *uhd* geeni *A. niger* –homeessa, saatiin Δ*gaaA udh* Δ*39114*-kanta, joka tuotti nestekasvatuksessa 4.3 ± 0.2 g/l galaktaarihappoa galakturonihaposta. Kiintoaine- ja nestekasvatuksessa käytettäessä sitrushedelmien kuorijätettä substraattina kanta tuotti 15.4 ± 0.6 ja 82.3 ± 3.9 mg galaktaarihappoa/g sitrushedelmien kuorijäte kuivapainona.

Avainsanat D-galakturonihappo, galaktaarihappo, sitruskuoret, filamenttihome, *Aspergillus niger*, CRISPR/Cas9, homologinen rekombinaatio

Preface

This Master's thesis was conducted at VTT Technical Research Centre of Finland Ltd in the Metabolic Engineering team from September 2015 to May 2016.

I would like to thank my thesis supervisor Associate Professor Alexander Frey for giving comments and feedback on my thesis. Your feedback had helped me during the writing process.

I would like to express my gratitude to my thesis advisor Docent Peter Richard for giving me this opportunity to work in the field related to genetic engineering. Peter's enthusiasm towards science and positive attitude has motivated and inspired me. Thank you for your support and guidance. I also wish to thank Joosu Kuivanen who has taught and helped me with the practical experiments. It has been a pleasure to work in the yeast-mould lab at VTT. Thanks to everyone in the lab for the friendliness and the guidance to the working model in the lab.

Finally, I would like to thank my family and friends for the support, especially my sister Daphne who is always asking about the conditions of my moulds. I also wish to give specials thanks to Miika Karjomaa who helped proofread my text.

Espoo, May 23rd 2016

Jasmin Wang

Contents

Abbreviations

5-FOA	5-fluoro-orotic acid
AP	Apiogalacturonan
Cas	CRISPR-assosiated enzyme
CPW	Citrus processing waste
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
dCas9	Dead Cas9
DM	Dry mass
DSB	Double-strand break
FDCA	Furan-2,5-dicarboxylic acid
GalA	Galacturonic acid
GalA	D-galacturonic acid
GlcA	Glucaric acid
gRNA	Guide RNA
HG	Homogalacturonan
HPLC	High pressure liquid chromatography
HR	Homologous recombination pathway
JGI	Joint Genome Institute
McA	mucic acid, galactaric acid
McA	Galactaric acid, mucic acid
NHEJ	Non-homologous end joining pathway
ORF	Open reading frame
PAM	Protospacer adjacent motif
PCR	Polymerase chain reaction
PEF	Polyethylene furanoate
PEG	Polyethylene glycol
PET	Polyethylene terephthalate
PVB	Polyvinyl butyral
PVC	Polyvinyl chloride
RG-I	Rhamnogalacturonan I
RG-II	Rhamnogalacturonan II
RVD	Repeat variable di-residue
RVD	Repeat variable di-residue
SF	Surface fermentation
SmF	Submerged fermentation
SmSF	Semisolid-state fermentation
SSB	Single-strand break
SSF	Solid-state fermentation
SSF	Solids-state fermentation
TALE	Transcription activator-like effector
TALEN	Transcription activator-like effector nuclease
tracrRNA	Transactivating CRISPR RNA
XGA	Xylogalacturonan

1 Introduction

For decades, microorganisms have been exploited for the production of food ingredients. Along with the development of genetic engineering, the utilization of organisms is no longer limited only to the food industry. Genetic engineering has enabled the modification of organisms to produce the desired product with higher titer and even express genes derived from another species. *Aspergillus niger* is one of the most used filamentous fungi in industrial biotechnology and it has been used to produce certain organic acids, such as gluconic and citric acid. Citric acid is a widely used product in different industries. Moreover, *A. niger* can also produce various enzymes, such as pectinases and amylases.

Hexaric and hexuronic acids are sugar acids derived by oxidizing a sugar. The molecular formula for hexaric acid is $C_6H_{10}O_8$ and for hexuronic acid is $C_6H_{10}O_7$. Hexaric acids have a carboxyl group both at C_1 and C_6, whereas hexuronic acids have an aldehyde group at C_1 and carboxyl group at C_6. Examples of hexaric and hexuronic acid are galactaric acid and D-galacturonic acid, respectively (Figure 1).

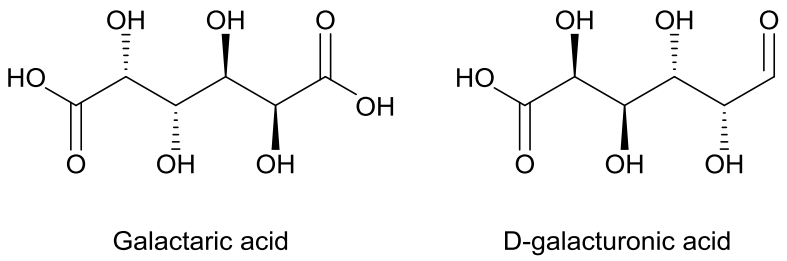

Figure 1 Structural formula of galactaric acid and acyclic D-galacturonic acid.

D-galacturonic acid is one of the main constituent of pectin, which can be obtained from inexpensive citrus peel and sugar beet pulp. Galactaric acid, on the other hand, is an interesting material, because it and its derivatives have many potential applications, especially in polymer synthesis. The filamentous fungus *A. niger* can be engineered to produce galactaric acid from galacturonic acid, but the mould itself catabolizes galactaric acid lowering the yield (Mojzita *et al.,* 2010). Therefore, the production of galactaric can

be improved by disrupting the galactaric acid catabolism, but the pathway in *A. niger* is still unknown. The ability to utilize mould in the production of value-added products directly from inexpensive raw materials such as pectin would be advantageous.

The objectives of this Master's thesis were to unravel the pathway for galactaric acid catabolism in *A. niger* by targeted deletion of the genes which are upregulated on galactaric acid, characterize the proteins encoded by these genes, characterize mutant strains and construct a galactaric acid producing strain. *A. niger* strains were engineered using homologous recombination and CRISPR/Cas9 genome editing tool.

LITERATURE REVIEW

2 *Aspergillus niger*

Aspergilli are widely used for the production of food ingredients, enzymes and antibiotics. They are tolerant towards acids and have a broad substrate spectrum. Two industrially important Aspergilli are *Aspergillus niger* and A*spergillus oryzae*. They both have GRAS (generally recognized as safe) status, which is an advantage when planning and designing new processes and products. For example, new fermentation processes and products utilizing these fungi are more easily approved by the authorities than processes utilizing organism without GRAS-status. *A. niger* has been used in enzyme and organic acid production. (Knuf and Nielsen, 2012) It is also used for biotransformations and waste treatment (Schuster *et al.*, 2002). *A. oryzae* has been used in food industry for sake, soy sauce and miso production for several centuries in Japan. (Knuf and Nielsen, 2012) In this chapter the focus is going to be on *A. niger*.

A. niger is composed of black-spored Aspergilli. It is a haploid, asexual filamentous fungus (Mable and Otto, 1998; Meyer *et al.*, 2007). *A. niger* reproduces by forming spores called conidia, which are uninucleates (Nga *et al.*, 1975). It grows aerobically on organic matter and can be found in soil, compost and on decaying plant material. *A. niger* is able to grow in the wide temperature range (6 – 47 °C) and pH range (pH 1.4 – 9.8). Its optimum growth temperature is at 35 – 37 °C. (Schuster *et al.*, 2002)

2.1 Industrial applications

In industry, *A. niger* has been used for citric acid, gluconic acid and industrial enzyme production. Another interesting application is using the fungus in waste treatment.

In 1919, citric acid was produced for the first time by fermentation with *A. niger*. The wild-type *A. niger* ATCC 1015 is capable of producing citric acid (Andersen *et al.*, 2011). Citric acid is mainly used in food and beverage industries as a primary acidifying agent, for example, in soft drinks, desserts and wine. (Schuster *et al.*, 2002) It is also used in cosmetics, in the pharmaceutical industry as antioxidants, in the detergent industry as a

phosphate substitute, and in the chemical industry as a foaming agent (Max *et al.*, 2010). In 2007, the annual production of citric acid was over 1.6 million tons and its demand is still increasing due to its numerous applications and the potential usage of citric acid in biopolymers, for drug delivery, tissue engineering for culturing a variety of cells and other biomedical applications (Dhillion *et al.*, 2010; Knuf and Nielsen, 2012).

Citric acid is mainly produced by submerged fermentation (SmF) with *A. niger* using a media containing glucose or sucrose. However, small- and medium-scale industries use surface fermentation (SF) for citric acid production, because the operation is easier, energy costs are lower and only simple equipment is needed. In the SF, *A. niger* grows as a thick floating mycelia mat on the surface of the growth media trays, which are on the shelves of the fermentation chambers. The SF is not suited for large-scale production, because the production would be unprofitable. It requires a large space and is time-consuming. Furthermore, it is vulnerable to contaminations and a large amount of heat is generated. The SmF requires lower investment and maintenance costs compared to the SF. (Max *et al.*, 2010; Dhillion *et al.*, 2010) Recently, there is an increased interest of using solid-state fermentation (SSF) for citric acid production. In the SSF, microorganisms grow on solid materials without the presence of free liquid. The SSF has many advantages over the SmF, such as solid waste management, reduced wastewater production, wide range of low-cost media, higher yield, production of high value products and lower energy requirements. (Dhillion *et al.*, 2010; Torrado *et al.*, 2011) In addition, in the SSF aeration is higher and low water activity reduces the risk of bacterial contamination (Torrado *et al.*, 2011). On the other hand, scale-up and controlling the process conditions are difficult in the SSF compared to the SmF (Dhillion *et al.*, 2010). Fermentors, such as, Erlenmeyer conical flasks, glass incubators, trays, rotating and horizontal drum bioreactors, packed bed column bioreactor, single-layer packed bed and multilayer packed bed have been used for citric acid production using SSF. The SSF process still requires improvements in several areas, for example yield and automation of the process, in order to increase its feasibility for industrial production. (Dhillion *et al.*, 2010).

Similar to citric acid, gluconic acid is also used in the food, beverage and pharmaceutical industries as an acidifying, flavour-enhancing and leavening agent (Roukas, 2000; Ramachandran *et al.*, 2006). Cement, textile and chemical industries also utilize gluconic

acid (Singh *et al.*, 2003). Annually, 60 000 tonnes of gluconic acid is produced and of its salt derivatives sodium gluconate is the most common. Sodium gluconate can chelate calcium and other di- and trivalent metal ions. Other gluconic acid derivatives are glucono-δ-lactone, calcium gluconate and iron gluconate. Glucono-δ-lactone is used in dairy industry for cheese curd formation and improving heat stability of milk. Calcium gluconate is used in calcium therapy for treating calcium deficiency. Similarly, iron gluconate is used in iron therapy. For example, zinc gluconate is used for treating the common cold and diseases caused by zinc deficiency. (Ramachandran *et al.*, 2006)

Gluconic acid is mainly produced by SmF with *A. niger*, in which D-glucose is oxidized to glucono-δ-lactone by glucose oxidase. Glucono-δ-lactone is then hydrolysed spontaneously near neutral pH or by lactonase to gluconic acid. Oxygen availability and pH are key parameters in gluconic acid production, since glucose oxidase needs oxygen for the oxidation of glucose and accumulation of organic acids produced by *A. niger* depends on the pH. If the pH falls below 3.5, the citric acid cycle will be triggered facilitating citric acid formation. Optimum pH for gluconic acid production is 5.5. (Ramachandran *et al.*, 2006). Singh *et al.* (2003) have investigated the production of gluconic acid using SmF, semisolid-sate fermentation (SmSF), SF and SSF with *A. niger* and varying glucose concentrations (40, 80, 120, 200 and 220 g/l). The gluconic acid production was the highest in all fermentation types when the initial glucose concentration was 120 g/l. Glucose utilization was most efficient in SmSF (89-97 %) and glucose up-take was higher in the SmF and the SmSF than in the SF and the SSF. Nevertheless, the conversion of glucose and gluconic acid yield was better in the SF and the SSF. The highest gluconic acid yield 94.7 % and concentration 106.5 g/l was obtained with the SSF. Similar to citric acid production, SSF is a promising fermentation type for gluconic acid production.

In the past years, there has been interest in using agricultural products and residues, such as, fig (the fruit of fig tree *Figus carica*), citrus and kiwi fruit peel and apple and grape pomace for citric acid and gluconic acid production utilizing *A. niger* (Roukas, 2000). For example, with orange peel Rivas *et al.* (2008) were able to obtain 9.2 g/l of citric acid by SmF. Furthermore, Torrado *et al.* (2011) were able to achieve higher citric acid yield with SSF (193.2 mg/ g dry orange peel) than with SmF (73.6 mg/g dry orange pel) from orange peel. When fig was used as the substrate for SSF the obtained citric

acid and gluconic acid yield was 64 g/kg dry fig and 490 g/kg dry fig respectively (Roukas, 2000).

In addition to organic acids, *A. niger* is also efficient in producing high amounts of hydrolysing enzymes, such as amylases, pectinases, hemicellulases and xylanases. For example, *A. niger* CBS 513.88 is an industrial enzyme-producing strain, which was derived from *A. niger* NRRL 3122 developed for glucoamylase A production. (Andersen *et al.*, 2011) The enzymes produced by *A. niger* have wide industrial applications, including glucoamylases used in the glucose syrup and alcohol industry and different pectinases, such as pectin esterases, endo- and exopolygalacturonidases and pectin lyases, used in wine and fruit juice production. Furthermore, the hemicellulases from *A. niger* can be added to dough, improving the doughs properties. (Schuster *et al.*, 2002)

Besides using *A. niger* as a production organism, the fungus has been used in waste treatment to remove metals from the environment. It has been listed as one of the top metal biosorbents. *A. niger* is able to adsorb metals to the cell wall components. Moreover, organic acids produced by the fungus form complexes with metals and leach or precipitate metals from materials. (Price *et al.*, 2001) Price *et al*. (2001) have studied six different fungi for their ability to remove copper and zinc from swine wastewater. Their results indicated that *A. niger* is the most efficient fungi for removing copper, thus further experiment with zinc absorption were conducted with *A. niger* only. It is known that *A. niger* possess a copper metallothionein, but the protein or the gene has not been sequenced. Metallothioneins are metal binding proteins which have been suggested to detoxify metals. Both living and dead *A. niger* were able to remove copper, but living fungus absorbed five times more copper on a mass basis. *A. niger* was capable of removing 91 % of the copper and 70 % of the zinc from treated swine effluent in 24 h.

Another potential waste treatment application utilizes *A. niger* to remove coloured dye substances in textile wastewater. *A. niger* is capable of adsorbing dyes and possibly also biodegrading them. (Assadi and Jahangiri, 2001; Omar, 2016) Assadi and Jahangiri (2001) have reported that *A. niger* could decolorize 97 % of textile wastewater in 20 h in airlift fermenter. According to Omar (2016), *A. niger* showed a strong ability to decolorize various reactive dyes (83-99 %) and grew well in a high dye concentration (500 mg/l). Carbon source and acidic conditions improved the decolorization in both reports.

3 Pectin-rich residues as a raw material

3.1 Structure of pectin

Pectin is structurally the most complex family of polysaccharides in nature and it can be found in plant cell walls. Pectin consist of a α-1,4-linked D-galacturonic acid backbone with sugar side chains. Approximately 70 % of pectin mass consists of galacturonic acid. There are five different types of pectin classified based on their structure: homogalacturonan (HG), xylogalacturonan (XGA), apiogalacturonan (AP), rhamnogalacturonan I (RG-I) and rhamnogalacturonan II (RG-II). HG is the most abundant and the simplest pectin since it is a linear galacturonic acid chain, which is partially methylesterified and acetylated. XGA is a HG substituted with xylose and AP is a HG substituted with D-apiofuranose. RG-II is the most structurally complex pectin. (Mohnen, 2008) It consists of a HG backbone with side chains consisting of 12 different sugars in over 20 different linkages. The sugars are D-glucuronic acid, L-rhamnose, D-galactose, L-arabinose, L-fucose, D-apiose, L-aceric acid, 2-*O*-methyl L-fucose, 2-*O*-methyl D-xylose, L-galactose, 2-keto-3-deoxy-D-*lyxo*-heptulosaric acid and 2-keto-3-deoxy-D-*manno*-octulosonic acid. (O'Neill *et al.*, 2004) In contrast to other pectin types, the RG-I backbone consists of a α-1,4-linked D-galacturonic acid and α-1,2-linked L-rhamnose repeating disaccharide unit with side chains consisting of various amount of L-arabinose and D-galactose. (Mohnen, 2008)

3.2 Galacturonic acid pathways

The main constituent of pectin, D-galacturonic acid, can be catabolized by some microorganisms. For example, the filamentous fungus *A. niger* is naturally capable of hydrolyzing pectin and catabolizing D-galacturonic acid to pyruvate and glycerol (Figure 2). In this reductive pathway, D-galacturonate (salt of galacturonic acid) is first reduced to L-galactonate by D-galacturonate reductase and then L-galactonate dehydratase removes a water molecule generating 3-deoxy-L-*threo*-hex-2-ulosonate. Next 2-keto-3-deoxy-L-galactonate aldolase splits 3-deoxy-L-*threo*-hex-2-ulosonate to pyruvate and L-glyceraldehyde. Finally, L-glyceraldehyde is reduced to glycerol by glyceraldehyde reductase. (Richard and Hilditch, 2009)

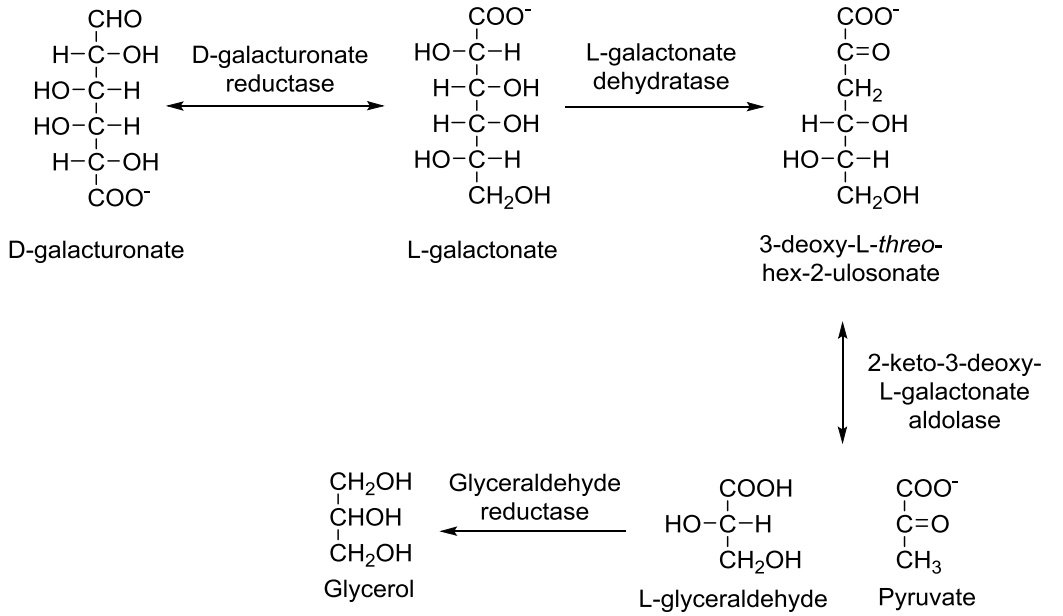

Figure 2 The fungal pathway for D-galacturonic acid catabolism, for example, in *Aspergillus niger*. D-galacturonate is converted to pyruvate and L-glyceraldehyde, which is further reduced to glycerol. (Richard and Hilditch, 2009, modified)

Other microorganisms such as *Agrobacterium tumefaciens* and *Pseudomonas syringae* can also catabolize D-galacturonic acid through an oxidative pathway (Figure 3). In the oxidative pathway, D-galacturonate is oxidized by D-galacturonate dehydrogenase forming D-galactaro-1,5-lactone, which rearranges in solution to more stable D-galactaro-1,4-lactone (Boer *et al.*, 2010; Parkkinen *et al.*, 2011). D-galactaro-1,5-lactone can also be converted to D-galactaro-1,4-lactone by D-galactarolactone isomerase (Bouvier *et al.*, 2014). *A. tumefaciens* has a galactarolactone cycloisomerase which can convert D-galactaro-1,4-lactone directly to 3-deoxy-2-keto-L-*threo*-hexarate (Andberg *et al.*, 2012). Previously it has been assumed that *g*alactaro-1,4-lactone can hydrolyses spontaneously at an alkaline pH or with the aid of lactonase to galactarate (Boer *et al.*, 2010). A water molecule is then removed from galactarate by galactarate dehydratase in a dehydration reaction forming 3-deoxy-2-keto-L-*threo*-hexarate or 3-deoxy-2-keto-D-*threo*-hexarate. However, it is not known which isomer is produced in the oxidative pathway. (Andberg *et al.*, 2012) It is still unclear whether *A. tumefaciens* or *P. syringae* also catabolize galacturonate via galactarate. Two types of galactarate dehydratases forming L-*threo*- and D-*threo*-hexarates have been discovered.

Dehydratase TalrD/GalrD from *Salmonella typhimurium* converts galactarate to 3-deoxy-2-keto-L-*threo*-hexarate (Wen *et al.*, 2007). GalrD-II from *Oceanobacillus iheyensis* (Rakus *et al.*, 2009) and GarlD-III from *A. tumefaciens* (Groninger-Poe *et al.*, 2014) convert galactarate to 3-deoxy-2-keto-D-*threo*-hexarate. In the following steps of galacturonic acid catabolism, 5-dehydro-4-deoxy-glucarate dehydratase dehydratates and decarboxylates 3-deoxy-2-keto-L-*threo*-hexarate to 2-ketoglutarate semialdehyde. The semialdehyde is then subsequently oxidized to 2-keto-glutarate by 2,5-dioxovalerate dehydrogenase. (Richard and Hiditch, 2009)

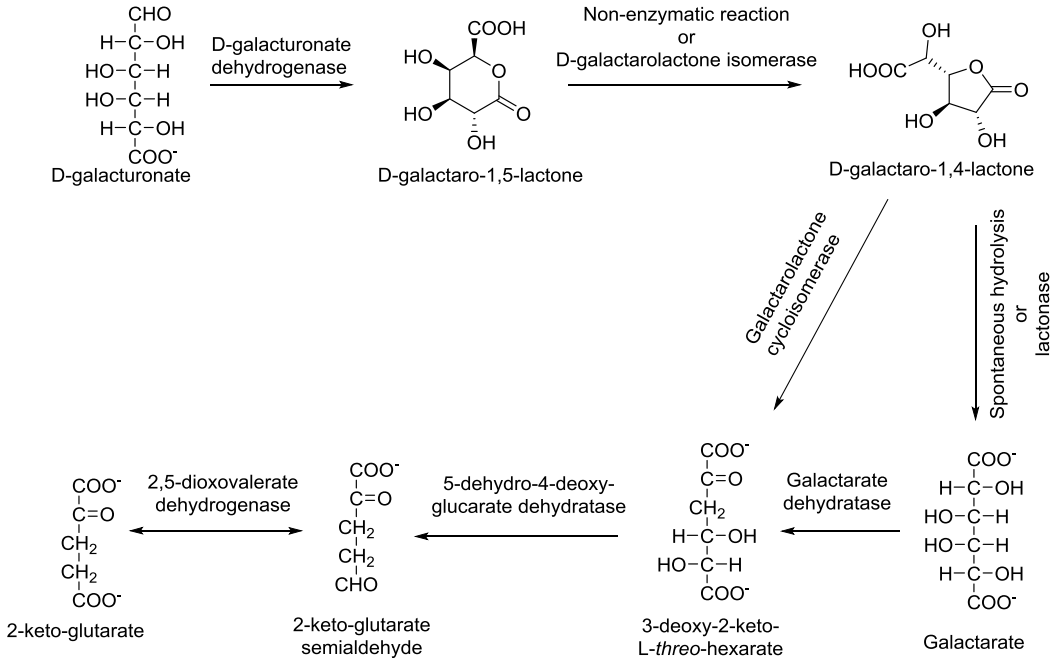

Figure 3 The oxidative pathway of D-galacturonic acid catabolism. In this pathway, D-galacturonate is oxidized to galactaro-1,5-lactone which can be enzymatically or non-enzymatically rearranged to galactaro-1,4-lactone. Galactaro-1,4-lactone then hydrolyses spontaneously or with the aid of lactonase to 3-deoxy-2-keto-L-*threo*-hexarate. However, in *Agrobacterium tumefaciens* galactaro-1,4-lactone is converted to 3-deoxy-2-keto-L-*threo*-hexarate.

E. coli catabolizes galacturonic acid through an isomerase pathway (Figure 4). In *E. coli*, D-galacturonate is converted to pyruvate and D-glyceraldehyde-3-phosphate in five steps. First, uronate isomerase converts galacturonate to D-tagaturonate. Next, D-tagaturonate reductase reduces D-tagaturonate to D-altronate, which is further converted to 3-doexy-D-*erythro*-hex-2-ulosonate by D-altronate dehydratase. Then

2-keto-3-deoxy-D-gluconate kinase phosphorylates 3-deoxy-D-*erythro*-hex-2-ulosonate to 3-deoxy-D-*erythro*-hex-2-ulosonate-6-phosphate, which is finally cleaved to pyruvate and D-glyceraldehyde-3-phosphate by 2-keto-3-deoxy-6-phosphogluconate aldolase. (Richard and Hilditch, 2009)

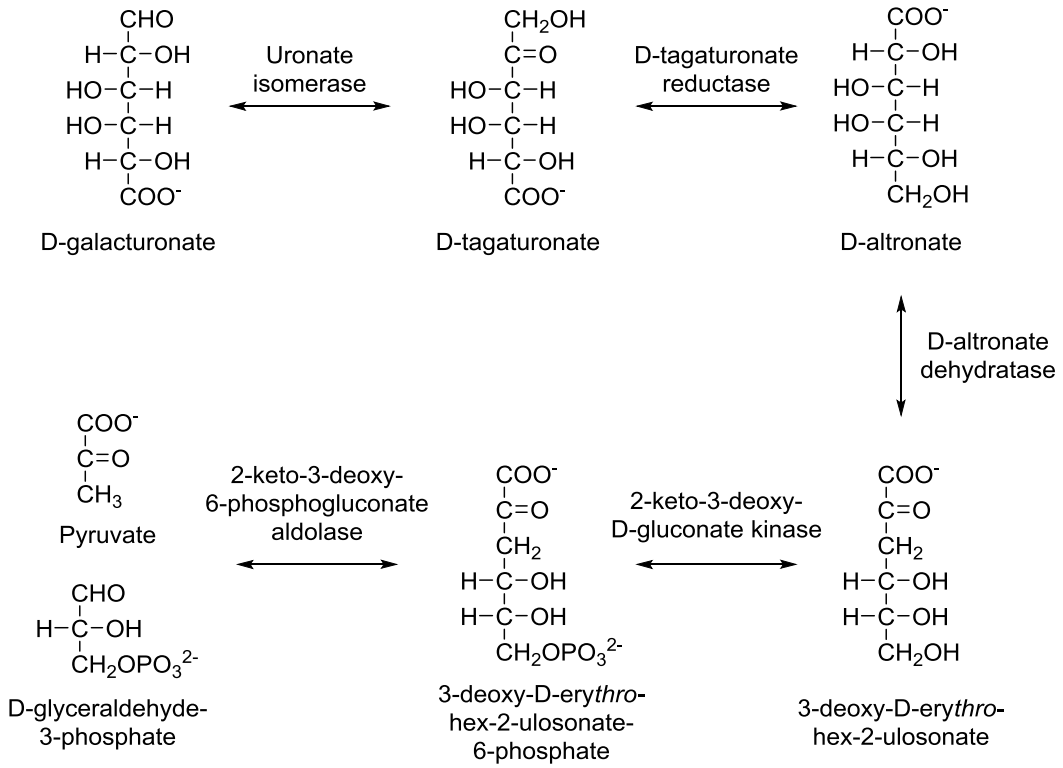

Figure 4 The isomerase pathway of D-galacturonic acid catabolism in *Escherichia coli*. In the pathway, D-galacturonate is converted to pyruvate and D-glyceralde-3-phosphate. (Richard and Hilditch, 2009, modified)

3.3 Applications of pectin-rich residues

Citrus juice production and sugar beet processing produce pectin-rich residues, which are mainly used as animal feed by the agricultural industry. Food industry also uses pectin produced from citrus peel as a gelling agent (Richard and Hilditch, 2009). Sugar beet pulp contains 24 % (Doran-Peterson *et al.*, 2008) and citrus processing waste (CPW) contains 25 % of pectin on a dry mass basis (Pourbafrani *et al.*, 2010). Processing the residues is energy consuming and the costs are high, because the residues require drying and pelletizing to prevent degradation (Richard and Hilditch, 2009). Instead of

processing the residues for applications such as animal feed, it would be more economical to find an alternative use.

Galacturonic acid in pectin-rich residues can be converted to useful compounds exploiting microorganisms. With recombinant bacteria, D-galacturonic acid can be converted to ethanol and galactaric acid. *E. coli* KO11, *Klebsiella oxytoca* P2 and *Erwinia chrysanthemi* EC 16 can produce ethanol with yields of 0.19, 0.12 and 0.16 g ethanol/g galacturonic acid, respectively. Using *E. coli* KO11 40 g/l ethanol could also be produced from sugar beet pulp treated with fungal enzymes in 24 h. (Doran *et al*., 2000) Recombinant *E. coli* can also be used to produce galactaric acid from sugar beet residue. The galactaric acid can then be further converted chemically to adipic acid by a deoxydehydration reaction catalysed by an oxorhenium complex reaction followed by a Pt/C-catalyzed hydrogenation. (Zhang *et al.*, 2016)

For ethanol production, *S. cerevisiae* would be an excellent organism, because it can tolerate high substrate and ethanol concentrations and low pH and oxygen levels. However, *S. cerevisiae* cannot naturally ferment galacturonic acid. (Edwards and Doran-Peterson, 2012) Therefore, it would need to be engineered by introducing a heterologous pathway. Furthermore, *S. cerevisiae* may not be able to assimilate galacturonic acid, hence, a plasma membrane galacturonic acid transporter should also be introduced. (van Maris *et al.*, 2006)

Galacturonic acid can also be converted to L-ascorbic acid (Kuivanen *et al.*, 2015) and galactaric acid (Mojzita *et al.,* 2010) by genetically engineered *A. niger*. Kuivanen *et al.* (2015) engineered *A. niger* to produce L-ascorbic acid (vitamin C) directly from galacturonic acid and CPW. First, the second enzyme L-galactonate dehydratase of the fungal galacturonic acid pathway was deleted. Then the heterologous enzymes L-galactono-1,4-lactone lactonase and L-galactono-1,4-lactone dehydrogenase from the plant L-ascorbic acid pathway were introduced into *A. niger* ATCC 1015 *ΔpyrG* strain. L-gulono-1,4-lactone lactonase from animal L-ascorbic acid pathway was also introduced into the *A. niger* instead of L-galactono-1,4-lactone lactonase. As a result, *A. niger* metabolized galacturonate to L-galactonate, L-galactono- 1,4-lactone and finally to L-ascorbic acid. Kuivanen *et al.* (2015) achieved a titre of 170 mg/l L-ascorbic acid. Production of galactaric acid by A. niger will be discussed in chapter 4.2.

4 Galactaric acid

4.1 Applications of galactaric acid and its derivatives

Galactaric acid, also known as *meso*-galactaric acid or mucic acid, is an aldaric acid, which is used as chelator in skin care products (Sauermann *et al.*, 2004), formerly in the1920s as leavening agent in self-rising flour and in dyeing textiles (Anonymous, 1922) and as a sequestering agent (Van Es *et al.*, 2015). It also has potential applications in polymer synthesis as a platform chemical, which are probably the most important applications.

Galactaric acid can be used to synthetize various types of polyhydroxypolyamides (hydroxylated nylons) with slightly different properties (Ogata *et al*., 1997; Kiely *et al.*, 2000). Galactaric acid can also be converted to several compounds, for example, adipic acid (Shiramizu and Toste, 2013; Li *et al.*, 2014), furan-2,5-dicarboxylic acid (FDCA) (Lewkoswski, 2001), 2,3,4,5-tetra-*O*-methylgalactaric acid (Mancera *et al.*, 2003), 2,3,4,5-tetra-*O*-acetylgalactaric acid (Mehtiö *et al.*, 2015), dibutyl 2,5-furandicarboxylate and dibutyl 2,3-furandicarboxylate (Taguchi *et al.*, 2008). All these compounds can be polymerized. Derivatives of galactaric acid have numerous applications. For example, adipic acid is primary used in the production of polyamide nylon-6,6. Other usages are in the production of polyurethane, as a plasticizer in polyvinyl chloride (PVC) and polyvinyl butyral (PVB) production and as an additive in cosmetics, gelatins, lubricants, fertilizers, adhesives, insecticides, paper and waxes. (Van de Vyver and Román-Leshkov, 2013) Similar to adipic acid, FDCA also has many applications. FDCA is an ingredient of fire foam with applications in both pharmacology and polymer synthesis (Lewkoswski, 2001). In pharmacology, it is utilized in the preparation of artificial veins for transplantation (Fraefel *et al.*, 2983). In polymer production, FDCA could replace widely used fossil based terephthalic acid (Wu *et al.*, 2014). Polyethylene furanoate (PEF) produced from FDCA has similar properties as polyethylene terephthalate (PET), making PEF a promising product with the potential to replace PET in making plastic bottles. (Anonymous, 2015; Liu B. *et al.*, 2015) This would be more sustainable since FDCA can be produces from bio-based materials.

4.2 Production of galactaric acid

Galactaric acid is currently produced commercially from D-galactose by oxidation with nitric acid (Acree, 1931), which is an expensive process and, moreover, produces toxic waste. Alternatively, chlorine could be used to oxidize galactose to galactaric acid with a nitroxide catalyst with the yield of 75-80 % (Nabyl *et al.*, 2002). Galactaric acid could also be produced from D-galacturonic acid by electrolytic oxidation with 80 % yield (Van Es *et al.*, 2015).

Inexpensive pectin-rich residues could provide a cost effective alternative for producing galactaric acid. A method of producing galactaric acid utilizing biotechnology is to genetically engineer filamentous fungi for the production. Mojzita *et al.* (2010) have genetically engineered *Trichoderma reesei* and *A. niger* for galactaric acid production by disrupting D-galacturonate catabolic pathway. They deleted a gene which encodes D-galacturonate reductase, *gar 1* in *T. reesei* and *gaaA* in *A. niger*, and expressed the bacterial gene *udh* encoding D-galacturonate dehydrogenase derived from *A. tumefaciens* in *T. reesei* and the codon-optimized *udh* gene in *A. niger* (Figure 5). D-galacturonate dehydrogenase converted galacturonic acid to galactaro-lactone which then spontaneously hydrolysed to galactaric acid. Mojzita *et al.* (2010) obtained the yield of 1.08 g galactarate/g galacturonate with *T. reesei* in 211 h, whereas with *A. niger* only 0.16 g galactarate/g galacturonate in 4 days. They suggested that *A. niger* catabolizes galactaric acid, resulting in a poor intracellular and extracellular accumulation of galactaric acid and thus a lowered yield. As a conclusion, in the case of *A. niger*, catabolism of galactaric acid should be disrupted in order to obtain better production yields. However, this catabolic pathway is still unknown. If galactaric acid can be produced by *A. niger*, the downstream processing would be simplified as galactaric acid precipitates at a low pH.

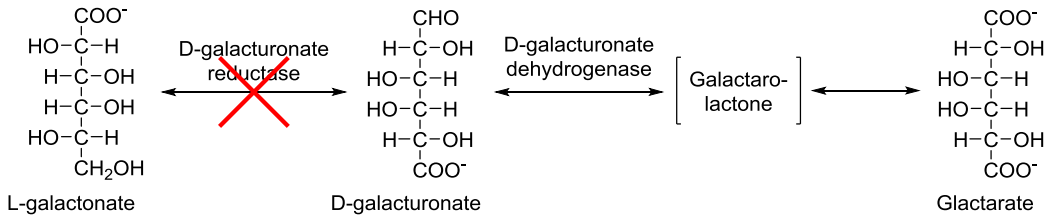

Figure 5 Production of galactaric acid with genetically engineered *Trichoderma reesei* and *Aspergillus niger* by disrupting D-galacturonic acid catabolism. Gene coding D-galacturonate reductase was deleted and a bacterial gene coding D-galacturonate dehydrogenase was expressed in both strains. (Mojzita *et al.*, 2010, modified)

Galactaric acid can also be produced by engineered recombinant *E. coli* from sugar beet residue (Zhang *et al.*, 2016). Overexpressing the *udh* from *A. tumefaciens* in the *E. coli* BL21(DE3), resulted in the conversion of 5, 10 and 20 g/l of pure D-galacturonic acid to 4.3, 8.3 and 11.5 g/l galactaric acid respectively in 48 h. The obtained yield was lower than the theoretical yield which was speculated to be due to the native catabolism of galacturonic and galactaric acid in *E. coli.* Therefore, to improve galactaric acid production the genes *uxaC* encoding uronoate isomerase in the galacturonic acid catabolic pathway and *garD* encoding D-galactarate dehydrogenase in the galactaric acid catabolic pathway were deleted from *E. coli* BL21(DE3). In a medium containing 10 g/l glucose, 10 g/l L-arabinose and 10 g/l D-galacturonic acid, the *E. coli* BL21(DE3) *udh* strain converted 68.5 % of the D-galacturonic acid to galactaric acid with a titer of 7.4 g/l and the *E. coli* BL21(DE3) Δ*uxaC* Δ*garD udh* strain converted 95.4 % of the D-galacturonic acid to galactaric acid with a titer of 10.3 g/l in 48 h. Sugar beet residue treated with a combination of cellulose, pectinase and Viscoayme® L in 50 mM citrate buffer at 50 °C produced a mixture of 22 % glucose, 17.4 % L-arabinose and 16.5 % D-galacturonic acid from the total residue. Using the supernatant of the enzyme treatment as the substrate, *E. coli* was able to convert 86 % of the D-galacturonic acid to galactaric acid in 48 h. The produced amount of galactaric acid was 6.9 g/l from 7.8 g/l of D-galacturonic acid. (Zhang *et al.*, 2016)

4.3 Catabolism of galactaric acid in microorganisms

E. coli and *A. tumefaciens* are capable of catabolizing galactaric acid. In *A tumefaciens*, galactaric acid may be formed as an intermediate in the galacturonic acid pathway when D-galacturonate is converted to galactarate in three steps. Catabolism of galactaric acid

in *A. tumefaciens* was described in the previous chapter (Figure 3). The first step of the galactaric acid catabolism in *E. coli* and *A. tumefaciens* is the conversion of galactarate to 3-deoxy-2-keto-hexarate by galactarate dehydrogenase. (Groninger-Poe *et al.*, 2014; Richard and Hilditch, 2009) Then the pathways differ, and in *E. coli* (Figure 6) 2-dehydro-2-deoxyglucarate aldolase splits 3-deoxy-2-keto-L-threo-hexarate to pyruvate and L-tartonate semialdehyde. In the following step, L-tartonate semialdehyde is reduced to D-glycerate by 2-hydroxy-3-oxopropionate reductase. (Richard and Hilditch, 2009)

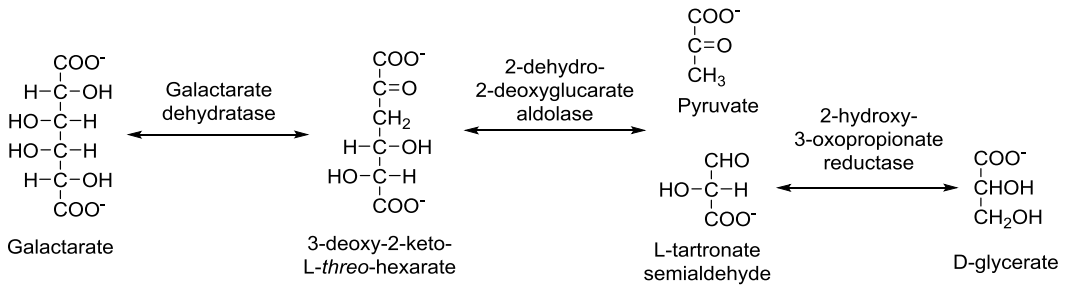

Figure 6 Galacturonic acid catabolic pathway in *Escherichia coli* (Richard and Hilditch, 2009). In this pathway, galactaric acid is converted to pyruvate and D-glycerate in three steps.

5 Genetic engineering of *Aspergillus niger*

Traditionally, strain improvements have been done using classical methods: random mutagenesis and screening. Nowadays, tools for genetic modifications have been developed, enabling more targeted gene modification and metabolic engineering. (Knuf and Nielsen, 2012) Methods and tools covered in this chapter are homologous recombination, CRISPR/Cas9 and TALENs.

Aspergilli host systems offer an alternative to *E. coli* and yeasts. They can perform post-translational modifications efficiently and express heterologous eukaryotic proteins in a correctly folded form. (Ward *et al.*, 2005) For example, *A. niger* can be a suitable host for heterologous protein expression (Punt *et al.*, 2002; Schuster *et al.*, 2002). On the other hand, transferring DNA into Aspergilli is extremely inefficient compared to *E. coli* and yeasts. Natural plasmids able to serve as a vector backbone for Aspergilli have not been

discovered. Therefore, expression vectors for recombinant protein are usually integrated into the chromosomal DNA and screened using different selection markers. (Fleißner and Dersch, 2010)

There are different transformation methods for introducing genes to Aspergilli: protoplast method, electroporation, *Agrobacterium* transformation and biolistic methods. Protoplast method is an efficient way and the most routinely used method to transform *A. niger.* In the protoplast transformation, the protoplasts are obtained by enzymatically digesting cell walls of germinated conidia or young mycelia. The protoplasts are incubated with plasmid DNA, polyethylene glycol (PEG) and in osmotic medium containing $CaCl_2$ and then spread on a selective solid medium. PEG is indispensable for transformation, because it helps the DNA to enter the cell. (Lakshmi Prabha and Punekar, 2004; Kawai *et al*., 2010; Meyer *et al.*, 2011) Electroporation and biolistic methods are less efficient than the protoplast method. Using *Agrobacterium* transformation high numbers of transformants have been obtained. (Fleißner and Dersch, 2010)

5.1 Homologous recombination

In homologous recombination, nucleotide sequences are exchanged between similar DNA molecules. It is a repair mechanism in eukaryotes used to repair double-strand breaks (DSBs) (Dudášová *et al.*, 2004). Homologous recombination is a commonly used method in the genetic engineering of fungi. For example, genes of *A. niger* can be deleted from the genome by replacing the target gene with a constructed deletion cassette (Figure 7). The deletion cassette consists of two flanking regions homologous to the sequences upstream and downstream of the target gene and, for example, a selection marker between these flanking regions.

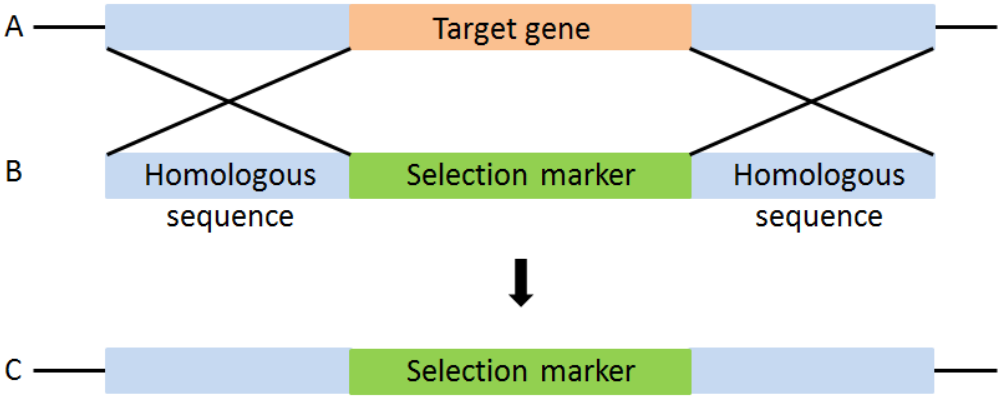

Figure 7 Gene deletion by homologous recombination in *A. niger.* (A) is part of the genome where the target gene is located. (B) is a deletion cassette designed for the target gene. When the deletion cassette is transformed to the cell, the cassette will replace the target gene and the gene is deleted forming (C).

There is a diverse selection of markers available for *A. niger* (Table 1): antibiotic resistance markers (*hph, oliC, bar*), auxotrophic markers (*pyrG, argB, niaD, trpC, sC*), and nutrional markers (*amdS, agaA*). From the above mentioned markers *pyrG, niaD, sC* and *amdS* are also counter selectable markers. (Lakshmi Prabha and Punekar, 2004; Meyer *et al.*, 2011; Dave *et al.*, 2012; Lenouvel *et al.*, 2002) Counter selection enables selection of both the presence and the absence of the marker. For example, $pyrG^{+}$ cells (cells harboring the *pyrG* gene) are able to grow in media lacking uracil or uridine, whereas $pyrG^{-}$ cells (cells without the *pyrG* gene) cannot grow in the same media. On the other hand, $pyrG^{+}$ cannot grow in media containing 5-fluoro-orotic acid (5-FOA), because it decarboxylates 5-FOA to toxic 5-fluorouracil. $pyrG^{-}$ does not have the ability to convert 5-FOA, hence it can grow in media containing 5-FOA. After the cassette has been constructed, it is transformed into *A. niger* and the gene of interest is deleted by homologous recombination. Using the same method, genes can also be integrated into the genome.

Table 1 Selection markers available for *A. niger* (Lakshmi Prabha and Punekar, 2004; Dave *et al.*, 2012).

Selection marker	Encoded function	Selection
hph	Hygromycin B phosphotransferase	Hygromycin resistance
oliC	Oligomycin-resistant mitochondrial ATP synthase	Oligomycin resistance
bar	Phosphinothricin acetyl transferase	Phosphinothricin resistance
pyrG	Orotidine-5'- phosphate decarboxylas	Uridine/uracil prototrophy, 5-fluoro-orotic acid resistance
argB	Ornithine carbamyltransferase	Arginine prototrophy
niaD	Nitrate reductase	Nitrate utilization, Chlorate resistance
trpC	Trifunctional enzyme of tryptophan biosynthesis	Tryptophan prototrophy
sC	ATP sulfurylase	Sulfate utilization, Selenate resistance
amdS	Acetamidase	Acetamide utilization, Fluoroacetamide resistance
agaA	Arginase	Arginine protoyrophy

The gene targeting frequencies in *A. niger,* and generally in filamentous fungi, are often extremely low compared to *S. cerevisiae* (Carvalho *et al.*, 2010). The cassette usually integrates randomly into the genome instead of the desired place, due to the non-homologous end joining pathway (NHEJ). NHEJ is the dominant DNA DSB repair mechanism in filamentous fungi compared to the homologous recombination pathway (HR). Deleting *kusA,* which encodes one of the proteins involved in NHEJ, improves homologous recombination. However, the strain is genetically unstable due to the increased sensitivity towards DNA damages caused by X-ray irradiation and UV exposure. (Meyer *et al.*, 2007) In addition, Meyer *et al.* (2007) have shown that the lengths of the flanks in the cassette affect the homologous recombination frequency.

With 1500 bp flanks, the homologous recombination frequency of *A. niger* Δ*pyrG*Δ*kusA* strain was as high as 98 %. Shorter flanks resulted in a notably lower frequency.

There is only a limited amount of selection markers available and therefore it would be advantageous to delete genes without leaving a marker on the genome. Delmas *et al.* (2014) have presented a method in which selection markers can be recycled for deleting genes in filamentous fungi. In this method, a plasmid with a selection marker is integrated to the genome and later excised. The method was demonstrated using filamentous fungi A*. niger* and *Talaeomyces versatilis* by deleting gene *xlnR* encoding xylanolytic transcriptional regulator. The plasmids used for recombination were non-replicative and harboured *pyrG* gene and flanking regions. When the plasmid was integrated to the genome and the *xlnR* gene was deleted, the cells were grown on a medium containing uridine to release the selective pressure for plasmid integration, resulting in the excision of the plasmid and the recovery of the *pyrG* selection marker. The excision using homology of the flanking region can lead to reversal to the wild-type locus or deletion of the target gene. The Δ*pyrG* strain was selected by plating to a medium containing uridine and 5-fluoro-orotic acid. Other available methods to perform deletions without leaving a selection marker on the genome are CRISPR/Cas9 (see 5.2) and TALENS (see 5.3).

Deleting an essential gene might result in formation of heterokaryons in primary transformants. Heterokaryon is a cell which has more than one nucleus and the nuclei are genetically different. Heterokaryotic strains are able to grow on selective conditions, but the two types of conidia produced from the heterokaryotic mycelium will not germinate on a selective medium. The conidia with untransformed nuclei will not grow because they lack the selection marker and the conidia with transformed nuclei will not grow due to the deletion of an essential gene. However, subculturing mycelia fragments is still possible. (Meyer *et al.*, 2007) Thus, if the deleted gene is essential for the survival of the cell, it would be difficult to obtain a pure culture. Southern blot analysis could be used for verifying heterokaryotic strains (Meyer *et al.*, 2007).

5.2 CRISPR/Cas9 system

Clustered regularly interspaced short palindromic repeats (CRISPR) is a bacterial and archaeal adaptive immune mechanism against foreign DNA derived from, for example, bacteriophages, plasmids and viruses. (Sander and Joung, 2014; Hsu *et al.*, 2014) CRISPR array contains short partially palindromic DNA repeats called CRISPR repeats that occur at regular intervals alternating with variable sequences called CRISPR spacers. The immune mechanism has three steps: adaption, CRISPR RNA (crRNA) biogenesis and targeting. In the adaption step, certain CRISPR-associated enzymes (Cas) acquire short exogenous DNA fragments which are inserted into the CIRSPR array as new spacers. Next, in crRNA biogenesis, CRISPR arrays are transcribed and processed into crRNAs. Biogenesis and processing of crRNA is unique for each CRISPR type. Finally, in the targeting step, crRNAs guide Cas nucleases for specific cleavage of homologous sequences and degrade the foreign DNA. (Barrangou and Marraffini, 2014; Hsu *et al.*, 2014)

There are three types of CRISPR system in which CRISPR/Cas9 is a type II CRISPR system. In the biogenesis of crRNA in type II CRISPR, a transactivating CRISPR RNA (tracrRNA) hybridized with CRISPR repeats forming a RNA duplex that is cleaved by endogenous RNase III liberating small crRNAs. The crRNA are further trimmed at their 5' end by an unknown nuclease. Cleavage of the DNA in type II CRISPR is carried out by endonuclease Cas9 guided by guide RNA (gRNA). Cas9 has two independent nuclease domains, HNH and RuvC, which both cleave one strand of the target DNA generating a blunt-ended DSB. (Barrangou and Marraffini, 2014; Hsu *et al.*, 2014)

In genome editing, Cas9 guided by synthetic gRNA creates a specific DNA DSB to the target site in the genome (Figure 8). The Cas9 used in the CRISPR/Cas9 genome editing is most commonly derived from *Streptococcus pyogenes* (Nødvig *et al.*, 2015), but Cas9 derived from *Streptococcus thermophilus* is also available (Hsu *et al.*, 2014). Recently a commonly used gRNA type is single gRNA, which is a fusion of crRNA and tracrRNA. Another available type of gRNA is the dual gRNA, in which crRNA and tracrRNA are expressed separately. In general, the results have been better with single gRNA than with dual gRNA. (Sander and Joung, 2014) The crRNA consist of a protospacer and CRISPR repeats. The protospacer consists of 20 nucleotides, which recognize the target site by base pairing. (Sander and Joung, 2014; Hsu *et al.*, 2014; Nødvig *et al.*, 2015) The

Cas9 and the gRNA can be expressed in the same vector *in vivo* (Nødvig *et al.*, 2015), but the gRNA can also be synthetized *in vitro* (Liu R. *et al.*, 2015).

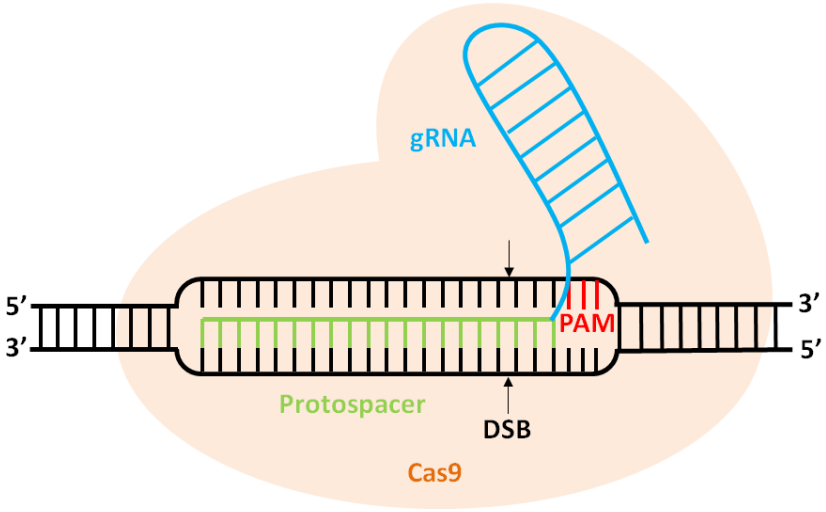

Figure 8 CRISPR/Cas9 system. The Cas9 endonuclease is guided by synthetic single guideRNA (gRNA) to the target site in the genome based on the protospacer sequence in the gRNA and protospacer adjacent motif (PAM) in the genome. Then Cas9 cleaves the DNA from 3-5 bp upstream of the PAM generating a double-stranded break (DSB).

Besides Cas9 and gRNA, the system also requires protospacer adjacent motif (PAM) to work efficiently. PAM is a three nucleotide sequence, which is adjacent to the binding site of gRNA in the downstream of the target site. Cas9 will cleave from the target site at 3-5 bp upstream of the PAM. For the Cas9 derived from *S. pyrogenes* the PAM sequence is NGG. There is also a less efficient alternative PAM sequence, NAG. (Sander and Joung, 2014; Nødvig *et al.*, 2015)

Cas9 can also be modified to expand its functionality. By inactivating the HNH or the RuvC nuclease domains of the Cas9 via point mutations, Cas9 can be modified to create single stranded break (SSB). These Cas9 nickases can be exploited for more specific NHEJ and HR, because SSBs are repaired by the high-fidelity base excision repair. Using a Cas9 nickase with two gRNAs and appropriately spaced target sites can mimic DSBs and mediate efficient indel formation. (Hsu *et al.*, 2014) A catalytically dead Cas9 (dCas9) can be utilized in controlling gene expression. dCas9 has mutations in both HNH and RuvC domains. It is a RNA guided DNA binding protein, which can be used to repress or

activate genes. Genes can be repressed by directing dCas9 to a promoter sequence interfering transcription initiation or to the template strand within an ORF blocking transcription elongation. Activation of genes can be done by fusing the dCas9 with different functional domains, for example, transcriptional activators and reporters to specific sites of the genome. In addition to altering the transcription states, chromatin states can also be monitored and even the three-dimensional organization of the genome can be rearranged. (Barrangou and Marraffini, 2014; Hsu *et al.*, 2014)

Recently, CRISPR/Cas9 technology has developed rapidly and become a new powerful tool in genome editing because it is very simple and easy to use. With the CRISPR/Cas9 technology it is possible to introduce multiple targeted DSBs to DNA triggering NHEJ pathways or HR to repair DSBs, and even insert a desired sequence to the targeted site. (Sander and Joung, 2014; Hsu *et al.*, 2014; Nødvig *et al.*, 2015) NHEJ-mediated repair can result in variable length of indels at the DSB site. Genes can be inserted to the DSB site by using a donor DNA to trigger homologous recombination. (Sander and Joung, 2014; Hsu *et al.*, 2014) Using the CRISPR/Cas9 system, gene deletion can be done without leaving a selection marker or foreign DNA in the genome.

CRISPR/Cas9 system has been used to engineer bacteria, yeasts, mammalian cells (eg. mice, rats, rabbits and human cells), insects, worms, fish and plants (eg. rice and wheat). (Sander and Joung, 2014; Nødvig *et al.*, 2015) In the case of filamentous fungi, *A. fumigatus* (Fuller *et al.*, 2015), *A. aculeatus, A. niger, A. nidulans, A. carbonarius, A. luchuensis, A. brasiliensis* (Nødvig *et al.*, 2015), *A. oryzae* (Katayama *et al.*, 2015), *T. reesei* (Liu R. *et al.*, 2015) and *Pyricularia oryzae* (Arazoe *et al.*, 2015b) have been engineered successfully utilizing the CRISPR/Cas9. As a genome editing tool, CRISPR/Cas9 system is more efficient and precise at targeting genes than merely using homologous recombination.

5.3 TALENs

Similar to the CRISPR/Cas9 system, transcription activator-like effector nucleases (TALENs) induce DNA DSBs at the target site that will trigger NHEJ or HR (Ochiai and Yamamoto, 2015). In addition, two TALEN pairs can also be used to delete or inverse large chromosomal segments by targeting to the same chromosome (Joung and Sander,

2013). TALENs are artificial fusion proteins composed of a customizable DNA binding domain fused to a nonspecific nuclease domain from the FokI restriction enzyme (Figure 9). The whole TALEN consists of N-terminal domain, TALE repeat domains, C-terminal domain and FokI nuclease domain. Two TALENs are needed to bind to the target site as a dimer for cleaving the DNA. (Joung and Sander, 2013) FokI is a type IIS restriction enzyme, which cleaves DNA at a defined distance away from their recognition sites. The TALENs bind to the target sites of DNA separated by spacer sequence (13-33 bp) facing one another and introduce DSB to the spacer. (Ochiai and Yamamoto, 2015)

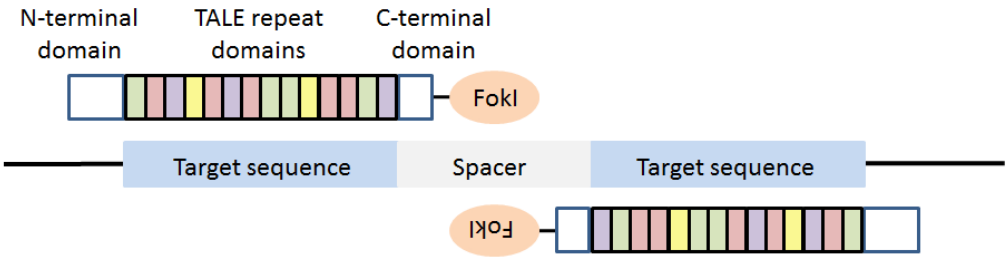

Figure 9 Transcription activator-like effector nuclease (TALEN) consists of DNA binding domain and an unspecific FokI endonuclease domain. N- and C-terminal domain and TALE repeat domains between terminal domains form the DNA binding domain. Each TALE repeats consists of 33-35 amino acid repeats, which have repeat variable di-residues (RVDs) at position 12 and 13. RVDs are responsible of recognizing the target sequence and binding to it. Two TALENs are needed for the cleavage of the DNA. The cleavage site is in the spacer region between the target sequences.

TALE repeat domains are composed of multiple nearly identical repeats (7-30) derived from transcription activator-like effectors (TALEs), which are proteins secreted by *Xanthomonas* spp. bacteria plant pathogen. These TALE repeats consists of 33-35 amino acid repeats. TALE repeat domains form an array that bind to single bases in the target sequence according to the repeat variable di-residues (RVDs) typically found at positions 12 and 13. RVDs are highly variable among the repeat unit and each RVDs NN, NI, HD and NG recognized DNA bases guanine, adenine, cytosine and thymine, respectively. There are also other more specific RVD for the DNA bases, but TALE repeat array using these RVDs show lower activity. The target sequence of TALENs begin with thymine, because the N-terminal domain recognizes the thymine adjacent to the nucleotide sequence recognized by the TALE repeat domains. TALENS also harbour non-repeat variable di-residue (non-RVD) at position 4 and 32, known as Platinum TALENS, which

have higher activity than TALENs without non-RVD variations. TALEs can be designed to bind to the specific target sequence by simply joining appropriate repeat units. (Joung and Sander, 2013; Ochiai and Yamamoto, 2015; Arazoe *et al.* 2015a) Truncation of the N- and C-terminal domains affect the DSB-introducing activities of TALENs. TALENs with N-terminal domain of +136 and C-terminal domain of +63 are the most commonly used and are highly active. (Ochiai and Yamamoto, 2015)

TALENs have been utilized in engineering for example, yeast, cricket, silkworm, fruitfly, zebrafish, plant *Arabidopsis thaliana* cows, pigs and human somatic and pluripotent stem cells (Joung and Sander, 2013). The only publication I found about engineering filamentous fungi with TALENs was reported by Arazoe *et al.* (2015a) in which they demonstrated TALEN-based genome editing via HR repair in filamentous fungus *Pyricularia oryzae* using Platinum-Fungal TALENs (PtFg TALENs). Arazoe *et al.* concluded that PtFg TALENs could improve the efficiency of targeted gene replacement by homologous recombination by up to 100%.

TALENs technology was introduced before the CRISPR/Cas technology. It seems to have higher specificity than CRISPR/Cas, but the preparation for genome editing of CRISPR/Cas is easier (Ochiai and Yamamoto, 2015) which makes the CRISPR/Cas the more preferable choice for genome editing.

EXPERIMENTAL STUDY

6 Materials and Methods

6.1 Microbial strains and media

A. niger, *S. cerevisiae* and *E. coli* strains used or constructed in this study are listed in Table 2. Gene *pyrG* encoding orotidine-5'- phosphate decarboxylase was used as a selection marker in all gene deletions of *A. niger*. *S. cerevisiae* ATCC 90845 was used for homologous recombination when constructing deletion cassettes and CEN.PK2-1D for expressing the open reading frame (ORF) of genes. *E. coli* TOP 10 was used for plasmid production. Liquid media and plates used in this study are listed in Table 3.

E. coli was cultivated in 5 ml Luria Broth (LB) medium with 100 µg/ml ampicillin for minipreps. *A. niger* cultivations for preparing protoplasts needed in transformation were performed in yeast Peptone medium with 3 % gelatin (YP-G). Mineral medium with 10 g/l galactaric acid (MM-McA), mineral medium with 10 g/l galactaric acid and 0.5 % xylose (MM-XMcA), yeast peptone medium with 10 g/l galactaric acid (YP-McA) and yeast peptone medium with 10 g/l D-glucaric acid (YP-GlcA) were used to examine the galactaric acid and glucaric acid consumption of the constructed mutant strains. Mineral medium with 20 g/l D-galacturonic acid (MM-GalA), mineral medium with 20 g/l D-galacturonic acid and 0.5 % xylose (MM-XGalA) and yeast peptone medium with 20 g/l D-galacturonic acid (YP-GalA) with pH 5 and 3 were used to investigate the production of galactaric acid of galactaric acid producing *A. niger* Δ*gaaA udh* Δ*39114* strain.

SCD-URA plates were used for screening *A. niger* and *S. cerevisiae* transformants. *A. niger* spores were produced on potato dextrose (PD) agar plates (incubation at 28 °C, 4-5 days) and collected with 2-4 ml spore stock solution (20% glycerol, 0.8% NaCl and 0.025 Tween 20) as suspension. Spores were filtered through a pipet tip with cotton wool.

S. cerevisiae was grown on YP plates. YPD medium was used to cultivate *S. cerevisiae* for transformation. For expressing ORFs and testing enzyme activities using enzyme activity assay, *S. cerevisiae* was cultivated in Y-min medium with SC-URA and 2 % glucose.

Table 2 Microbial strains used in this study. Gene *pyrG* encodes orotidine-5'- phosphate decarboxylase, *kusA* is a gene responsible for non-homologous end joining, *gaaA* encodes D-galacturonic acid reductase, *gaaC* encodes 2-keto-3-deoxy-L-galactonate aldolase and *udh* encodes D-galacturonate dehydrogenase.

Name	Parental strain	Genetic modifications	Reference
Aspergillus niger			
Wild type, ATCC 1015			Andersen *et al.*, 2011
Δ*pyrG*	ATCC 1015	Deletion of *pyrG* gene	Kuivanen *et al.*, 2012
Δ*pyrG*Δ*kusA*	Δ*pyrG*	Deletion of *pyrG* and *kusA*	Unpublished
Δ*39114*	Δ*pyrG*	Deletion of gene *39114*	This study
Δ*1090836*	Δ*pyrG*	Deletion of gene *1090836*	This study
Δ*1117792*	Δ*pyrG*	Deletion of gene *1117792*	This study
Δ*1141260*	Δ*pyrG*	Deletion of gene *1141260*	This study
Δ*1121140*	Δ*pyrG*	Deletion of gene *1121140*	This study
Δ*1162477*	Δ*pyrG*	Deletion of gene *1162477*	This study
Δ*1146483*	Δ*pyrG*	Deletion of gene *1146483*	This study
Δ*1170646*	Δ*pyrG*	Deletion of gene *1170646*	This study
Δ*39114*Δ*gaaA/C::udh*	Δ*39114*	Deletion of *gaaA* and *gaaC* and integration of *udh* derived from *P. syringae* to *gaa/C* locus	This study
Δ*39114*Δ*gaaA/C*	Δ*39114*	Double deletion of *gaaA* and *gaaC*	This study
Δ*39114*Δ*gaaA/C udh*	Δ*39114* Δ*gaaA/C*	Double deletion of *gaaA* and *gaaC* and random integration of *udh* derived from *P. syringae*	This study
Δ*gaaA udh*		Deletion of *gaaA* and integration of codon-optimized *udh* derived from *A. tumefaciens*	Mojzita *et al.*, 2010
Δ*gaaA udh* Δ*39114*	Δ*gaaA udh*	Deletion of gene *39114* from Δ*gaaA udh*	This study
Saccharomyces cerevisiae			
ATCC 90845			
CEN.PK2-1D		Expression of ORF of *A. niger* gene	
SC39114	CEN.PK2-1D	Expression of the ORF *39114* with pE39114 expression plasmid	This study
SC1090836	CEN.PK2-1D	Expression of the ORF *1090836* with pE39114 expression plasmid	This study
SC1121140	CEN.PK2-1D	Expression of the ORF *1121140* with pE39114 expression plasmid	This study
Escherichia coli			
TOP 10		Used for plasmid production	

Table 3 List of liquid media and plates used in this study for cultivation. ORF = open reading frame

Media	**Abbreviation**	**Usage**
Luria Broth	LB	*E. coli* cultivation
Yeast Peptone medium with 3 % gelatin, pH 7	YP-G	*A. niger* precultivation
Mineral medium with 1% galactaric acid, pH 7	MM-McA	*A. niger* cultivation with galactaric acid
Mineral medium with 1% galactaric acid and 0.5 % xylose, pH7	MM-XMcA	*A. niger* cultivation with galactaric acid
Yeast Peptone medium with 1% galactaric acid, pH 7	YP-McA	*A. niger* cultivation with galactaric acid
Yeast Peptone medium with 1% D-glucaric acid, pH 7	YP-GlcA	*A. niger* cultivation with glucaric acid
Mineral medium with 2% galacturonic acid, pH 5 or pH 3	MM-GalA	*A. niger* cultivation with galacturonic acid
Mineral medium with 2% galacturonic acid and 0.5 % xylose, pH 5	MM-XGalA	*A. niger* cultivation with galacturonic acid
Yeast Peptone medium with 2% galacturonic acid, pH 5 or pH 3	YP-GalA	*A. niger* cultivation with galacturonic acid
A. Nidulans defined minimal medium		*A. niger* cultivation with citrus processing waste
Yeast Peptone medium with 2% glucose	YPD	*S. cerevisiae* cultivation
Yeast nitrogen base without amino acids	Y-min	*S. cerevisiae* cultivation for ORF expression
Agar plates	**Abbreviation**	**Usage**
SCD-URA transformation plate		*A. niger* transformation using homologous recombination
Transformation plate with 400 µg/ml hygromycin (minimal medium)		*A. niger* transformation using CRISPR/Cas9 system
Synthetic complete medium without uracil containing 2 % glucose	SCD-URA	*A. niger* screening, *S. cerevisiae*
Synthetic complete medium without uracil containing 2 % D-galacturonic acid	SCGalA-URA	Verifying the deletions of *gaaA* and *gaaC* in *A. niger*
Potato dextrose agar	PD	*A. niger* spore germination
Yeast Peptone	YP	*S. cerevisiae*

6.2 Deletion of potential genes in galactaric acid pathway

6.2.1 Construction of deletion cassette

First, the plasmids containing deletion cassette were constructed. The cassette consisted of two flanking regions amplified from *A. niger* ATCC 1015 genome, an auxotrophic selection marker *pyrG* and a backbone plasmid. The deletion cassettes were designed to have a NotI restriction site in order to digest the cassette, which consist of flanking regions and the *pyrG*, from the plasmid. Flanking regions (1.5 kb) were amplified from the genomic DNA up- and downstream of the gene of interest. Geneious ver. 8.1.7 (Biomatters Ltd.) was used to design primers. The primers used for amplifying the flanking regions are listed in Appendix 1 Table 1. The backbone plasmid p2974 (pRS426, Colot *et al.*, 2006) linearized with EcoRI and BamHI has ampicillin as *E. coli* selection marker, bacteria origin of replicon *ori*, *URA3* gene as *S. cerevisiae* selection marker and yeast origin of replicon *2μ*. Then *S. cerevisiae* was transformed using Gietz protocol with (see 6.6.1) all these four fragments (two flanking regions, *pyrG* gene and the backbone plasmid) for homologous recombination using Gietz protocol to assemble the plasmid. Cells were plated on SCD-URA plates and incubated at 30 °C for 4 days.

Next, the grown *S. cerevisiae* transformants were collected from the plate and the plasmids were rescued from the transformants with a miniprep (QIAprep Spin Miniprep Kit). Then *E. coli* was transformed by electroporation with the plasmids and grown on LB plates containing ampicillin (100 μg/ml) at 37 °C overnight (electroporation protocol see 6.6.2). Colony PCR was performed to confirm that the *E. coli* transformants harboured the desired plasmid. The plasmids were then isolated from the confirmed *E. coli* colony. Finally, the plasmids were verified by the digestion with NotI and the sequencing of the 5' flanking region. The sequencing primer attached to the beginning of the *pyrG* gene and sequenced in the reverse direction.

6.2.2 Targeted deletion

The cassettes were digested with NotI from the plasmids containing the cassette and the solution containing the cassette was concentrated by ethanol precipitation, in which 1000 μl absolute ethanol and 40 μl 3 M sodium acetate were mixed with 300 μl cassette solution and incubated at -20 °C for 20 min. Then the mixture was centrifuged at 4 °C for 10 min at full speed. The DNA pellet was washed with 300 μl 70 % ethanol and

centrifuged at room temperature for 5 min at full speed. The pellet was dried and resuspended in 100 µl transformation buffer.

Deletion was performed by homologous recombination in *A. niger* by transforming the Δ*pyrG* (orotidine-5'- phosphate decarboxylase deletion strain) and Δ*pyrG*Δ*kusA* strains (orotidine-5'- phosphate decarboxylase and NHEJ deletion strain) with the cassettes (transformation see 6.6.3). Deletion was also performed using the CRISPR/Cas9 method, in which the Δ*pyrG* strain was transformed with the pFC-332 plasmid harboring Cas9 (Nødvig *et al.*, 2015), two single gRNAs and a donor DNA. The gRNAs were designed to guide Cas9 to cleave DNA both upstream and downstream of the gene to be deleted. GC content of the target sites were designed to be at least 40 %. The gRNAs were synthesized *in vitro* using GeneArt™ Precision gRNA Synthesis Kit (ThermoFisher Scientific). The oligonucleotides needed for the synthesis of gRNA are listed in Appendix 1 Table 5. The constructed deletions cassettes were used as donor DNA for the selection. Around 10 µg of gRNAs and deletion cassette and 1 µg of pFC-332 was used in transformation, except for deleting *1146483* only 5 µg of deletion cassette was used.

Transformants were screened by colony PCR using Phire Plant Direct PCR Kit (Thermo Scientific). The 5' flanking region was amplified with primers annealing to the beginning of the *pyrG* gene and outside of the 5' flanking region, the 3' flanking region with primers annealing to the end of the *pyrG* gene and outside of the 3' flanking region and part of the ORF was amplified with primers annealing internal to the gene of interest. The colonies were also screened by colony PCR using primers binding outside of the 5' flanking region and inside the ORF. The primers are listed in Appendix 1 Table 2. The screened colonies were also inoculated to SCD-URA masterplate (incubation 28 °C, 2 days) for further purification.

Colonies in which deletion had occurred were purified by single spore isolation. First, the colonies were inoculated from the SCD-URA masterplate to PD plate for sporulation (incubation 28 °C, 2 days). Then the spores were collected, mixed to spore stock solution and diluted 10, 100 and 1000 times with spore stock solution. 10 µl of diluted suspensions were plated to SCD-URA plate (incubation 28 °C, 2 days). Colonies on the SCD-URA plates were screened by colony PCR again and inoculated to PD plate. The strains went through single spores until a pure culture was obtained. Pure cultures were let to sporulate on PD plate, the spores were collected, and spore stocks were prepared.

Pure gluconokinase 1162477 mutants were also cultivated on 2 % gluconic acid plates as a screening method.

6.3 Identification and characterization of the proteins

6.3.1 Cultivation of deletion mutants with hexaric acid

The mutant strains were precultivated by inoculating 300-400 µl spores into 50 ml YP-G medium and incubating at 28 °C for 21 hours. Then the grown mycelia were collected by filtering through a sterile Glass Microfibre Filter (55mm diameter, Whatman, GE Healthcare Life Science, USA) with vacuum and washed with sterile water. The mycelia were resuspended in the MM-McA or YP-GlcA medium. Cultivation was done in 24 wells plates with two replicates. Around 0.02 g wet cells mass (in 0.5 ml MM-McA) of the mycelia were inoculated into 3.5 ml of MM-McA, MM-XMcA and YP-McA or YP-GlcA medium (Table 3), hence the initial biomass concentration was 5 g wet cell mass/l. Cultures were incubated at 28 °C, 800 rpm for 5 days. Samples of 0.5 ml were taken at 24 h, 48 h and 120 h and filtered with 24 wells 10 µm Melt Blown Polypro Filter (Whatman, GE Healthcare) or 96 well MultiScreen HTS DV Filter Plate (0.65 µm Hydrophilic Low Protein Binding, Durapore® Membrane).

6.3.2 HPLC analysis

The consumption of galactaric acid was analyzed with high performance liquid chromatography (HPLC) using a Animex Fast Acid column (100 mm by 7.8 mm, Bio-Rad Laboratories) linked to Animex HPX-87H organic acid analysis column (300 mm by 7.8 mm, Bio-Rad Laboratories). The columns were maintained at 55 °C and 5 mM H_2SO_4 was used as an eluent with flow rate 0.5 ml/min. Waters 2489 UV/Visible dual-wavelength UV (210 nm) detector was used for detection.

6.3.3 Expressing genes in *Saccharomyces cerevisiae*

First, the cDNA was generated to amplify the ORFs of the genes of interest for expression plasmid construction. The wild-type *A. niger* ATCC 1015 strain was cultivated in YP medium with 3% gelatin, collected and frozen in liquid nitrogen. Then total RNA was extracted using RNeasy Plant Mini Kit (Qiagen) and genomic DNA was removed using DNase I following Removal of Genomic DNA from RNA Preparations protocol

(Thermo Scientific). The cDNA was synthetized using Transcriptor High Fidelity cDNA Synthesis Kit (Roche) and the ORF was amplified using cDNA as the template. EcoRI restriction site and Kozak sequence were added upstream and BamHI restrictrion site downstream of the ORF. The primers for amplifying the ORFs are listed in Appendix 1 Table 3. The PCR product of the ORF was purified with a Miniprep column (Qiagen), digested with EcoRI and BamHI and purified by gel extraction. Then the ORF was ligated with the linearized expression vector p2159 by T4 ligase, and introduced into *E. coli*. The expression vector p2159 was derived from pYX212 (R&D systems) by modifying the multiple cloning site (Verho *et al.*, 2004) and digesting with EcoRI and BamHI. The whole ORF was sequenced. Sequencing primers are listed in Appendix 1 Table 3. Then the expression plasmids were transformed into *S. cerevisiae* CEN.PK-1D and the transformants were grown on SCD-URA plates. Two transformants were inoculated from the transformation plate to new SCD-URA plate to prepare glycerol stocks for further experiments.

6.3.4 Enzyme Assay

Crude extract for the enzyme assay was prepared by cultivating *S. cerevisiae* CEN.PK2-1D harbouring a gene expression plasmid in Y-min medium containing SC-URA and 2 % glucose at 30 °C, 250 rpm for overnight. The cells of overnight culture were collected (4 °C, 5000 rpm, 5 min), washed with cold sterile water and pelleted (4 °C, 5000 rpm, 5 min). The pellet was resuspended in 1 ml lysis buffer (0.1 M $NaPO_4$ containing complete protease inhibitor). The cells were broken with glass beads (diameter 0.5 mm) using the Precellys 24 (Bertin Technologies) for three times 30 s. The mixture was centrifuged (4 °C, 14 000 rpm, 20 min) and the supernatant also known as the crude extract was collected.

In the assay, galactaric acid was used as the substrate and the formation or consumption of NAD(P)H was followed at 340 nm using a spectrophotometer for 300 s with readings in 10 s interval. The 1 ml reaction mix contained $NaPO_4$ buffer (50 mM, pH 7), ATP (0.5 mM), $MgCl_2$ (1mM), galactaric acid (2mM), a cofactor (NADH or NADPH 0.2 mM, or NAD or NADP 0.5 mM), 10 µl crude extract. The activity of the 39114 enzyme was also tested using 10 mM glucaric acid as the substrate and NADH and NADPH as cofactors.

6.4 Construction of galactaric acid producing *A. niger* strains

6.4.1 Construction of the Δ*39114*Δ*gaaA/C*::*udh* and Δ*39114*Δ*gaaA/C udh* strains

The plasmid containing integration cassette for deleting the genes *gaaA* and *gaaC* (*gaaA/C*) and expressing *udh* in the Δ*39114* strain was designed to contain two flanking regions, *gpdA* promoter, *udh*, *trpC* terminator and *hph* coding regions and to be released from the plasmid with NotI. The used backbone plasmid p2974 had ampicillin resistance gene as *E. coli* selection marker, bacteria origin of replicon *ori*, *URA3* gene as the *S. cerevisiae* selection marker and yeast origin of replicon *2μ*.

Flanking regions were amplified from genomic DNA using primers oPEEL-053-oPEEL-056 (Appendix 1 Table 4). The gene *udh* with promoter, terminator and *hph* were amplified from the plasmid p6042 using primers oPEEL-051 and oPEEL-052 (Appendix 1 Table 4). Three different methods were used to assemble the fragments of the plasmid containing the integration cassette: homologous recombination in *S. cerevisiae*, Gibson Assembly and TOPO Cloning. In yeast homologous recombination the *S. cerevisiae* ATCC 90845 was transformed with the fragments. Then the resulting plasmid was rescued from the yeast and *E. coli* was transformed with the rescued plasmid by electroporation. In Gibson Assembly, Gibson Assembly Cloning Kit (BioLabs) was used to assemble the integration cassette. The fragments were assembled with the reaction mix provided by the kit and then *E. coli* was transformed with the plasmid by electroporation. In TOPO Cloning, the integration cassette was amplified as one fragment in one PCR reaction (Kapa HiFi PCR Kit) using previously amplified flanking regions and *udh* as the templates and oPEEL-053 and oPEEL-056 as the primers. The cassette was then purified by gel extraction and ligated with TOPO vector using Zero Blunt TOPO PCR Cloning Kit. Then *E. coli* was transformed with the ligated plasmid by electroporation for plasmid production. The *E. coli* transformants harboring the plasmids assembled using different methods were screened by colony PCR by amplifying the 5' and 3' flanking regions of the plasmid. The plasmid assembled using yeast homologous recombination was also digested with NotI.

The CRISPR/Cas9 method was also used to delete the *gaaA/C* gene from the Δ*39114* strain and integrate the *udh* gene into the same strain. Two gRNAs were designed to

guide Cas9 to cleave DNA at the *gaaA* and *gaaC* locus. The gRNAs were synthesized *in vitro* using GeneArt™ Precision gRNA Synthesis Kit and the oligonucleotides oPEEL-080 – oPEEL-083 (Appendix 1 Table 4). The *gpdA-udh* fragment (2.5 kb) used as a donor DNA, containing the *gpdA* promoter and the *udh* gene*,* was amplified using the plasmid p6042 as the template (Kapa HiFi PCR Kit). The primers oPEEL-084 and oPEEL-085 (Appendix 1 Table 4) used for the amplification were 80 bp long from which 60 bp was homologous to the genome. After all the different components needed for CRISPR/Cas9 were prepared, they were transformed into the Δ*39114* strain. Three different transformations were performed: 1) the Δ*39114* strain transformed with pFC-332 (as a control), 2) the Δ*39114* strain transformed with pFC-332 and two gRNAs and 3) the Δ*39114* strain transformed with pFC-332, two gRNAs and the *gpdA-udh* fragment. The transformation was performed with around 10 µg of gRNAs, 10 µg of donor DNA and 1 µg of the pFC-332 plasmid. The correct colonies were identified by colony PCR using primers oPEEL-086 and oPEEL-087 (Appendix 1 Table 4), which bind to the outside of the *gaaA/C* gene. Selection of the Δ*39114*Δ*gaaA/C* strain was also performed by cultivating on SC-URA plate with 2 % D-galacturonic acid (28 °C, 2 days). The Δ*39114* strain was used as a control strain.

6.4.2 Construction of the Δ*gaaA udh* Δ*39114* strain

The *39114* gene was deleted from *A. niger* Δ*gaaA udh* strain (Mojzita *et al.*, 2010) using the CRISPR/Cas9 method. In the Δ*gaaA udh* strain, the *pyrG* gene was used as a selection marker in the *gaaA* gene deletion and the *amds* gene in the codon optimized *udh* integration. Two *in vitro* gRNAs were designed to excise a DNA fragment (around 541 bp) of the ORF *39114* targeting the exon 2 and the exon 3. Three transformations with different conditions were performed: 1) the Δ*gaaA udh* strain with pFC-332 (as a control), 2) the Δ*gaaA udh* strain with pFC-332 and two gRNAs and 3) the Δ*gaaA udh* strain with pFC-332, gRNA cleaving from the exon 3 and the deletion cassette as the donor DNA. The transformations were performed with around 10 µg of the gRNAs, 10 µg of the donor DNA and 1 µg of the pFC-332 plasmid. The resulting Δ*gaaA udh* Δ*39114* transformants were screened by colony PCR using primers pairs oPEEL-001 and oPEEL-029, oPEEL-059 and oPEEL-060 and oPEEL-088 and oPEEL-089 (Appendix 1 Table 2).

6.5 Cultivation for galactaric acid production

6.5.1 Cultivation with galacturonic acid

The strains *ΔgaaA udh Δ39114* and *ΔgaaA udh* were grown on SCGalA-URA (SC-URA containing 2 % galacturonic acid) plate and also submerged cultivated with galacturonic acid at different scales: 4 ml and 50 ml. The wild-type strain ATCC 1015 was also submerged cultivated for comparison.

In the 4 ml-scale submerged cultivations, the strains *ΔgaaA udh Δ39114*, *ΔgaaA udh* and the wild-type strain were cultivated in two different mineral media, MM-GalA and MM-XGalA containing 20 g/l galactaric acid at initial pH 5 (Table 3). Cultivations were performed as in the cultivation of deletion mutants with galactaric acid (see 6.3.1) with few exceptions. In precultivation, the volume of the used spore suspension was 200 – 300 µl and the filtered mycelia were resuspended to MM-GalA medium. As earlier, there were two replicates, initial biomass concentration was 5 g wet cells mass/l and the cultures were incubated at 28 °C, 800 rpm for 5 days.

The 50 ml-scale submerged cultivations were performed in 250 ml Erlenmeyer flasks. The strains *ΔgaaA udh Δ39114*, *ΔgaaA udh* and the wild-type strain were cultivated in MM-GalA and YP-GalA with intitial pH 5 or 3 with three replicates. The media were inoculated with around 0.5 g mycelia resuspended to the same cultivation medium, so that the initial mycelia concentration was 10 g wet cell mass/l. Cultures were incubated at 28 °C, 200 rpm for 5 days.

The formation of galactaric acid and the consumption of galacturonic acid during the cultivations were followed by HPLC. The samples were taken at 24 h, 48 h, 72 h, 96 h and 120 h and prepared by filtering with 96 well MultiScreen HTS DV Filter Plate.

6.5.2 Cultivation with citrus processing waste

CPW prepared according to Kuivanen *et al.* (2014) was used as a substrate for SSF and SmF. The dry mass (DM) of the prepared CPW was determined by weighing the CPW before and after drying in an oven at 100 °C overnight. The fermentations were performed in 250 ml Erlenmeyer flasks with three replicates.

In the SSF, 6 ml of sterile water or 13 g/l $(NH_4)_2SO_4$ containing an inoculum of 2×10^7 *ΔgaaA udh Δ39114* spores were mixed with 2 g of CPW. The cultures were incubated at 28 °C for 137 h. After 3 days of cultivation, 2 ml of sterile water or $(NH_4)_2SO_4$ was added to the cultures. The fermentation products were extracted by adding 30 ml of sterile water to the cultures followed by an incubation at 28 °C, 200 rpm for 1 h. The suspension was filtered with Glass Microfibre Filter (55mm diameter, Whatman, GE Healthcare Life Science, USA) with vacuum and the supernatant was analyzed with HPLC.

In the SmF, 2 g of CPW was incubated with 50 ml *A. nidulans* defined minimal medium ($NaNO_3$ 6 g/l, KCl 0.52 g/l, $MgCl_2$ 0.52 g/l and KH_2PO_4 1.52 g/l, Barratt *et al.*, 1965) inoculated with 10 g wet cell mass/l of the wild-type, the Δ*gaaA udh* and the *ΔgaaA udh Δ39114* mycelia. The fermentations were done at 28 °C, 200 rpm for 5 days. Samples for HPLC was taken every 24 h and treated as in previous cultivations before analyzing with HPLC.

6.6 Transformation procedures

6.6.1 Transformation of *S.cerevisiae* – Gietz protocol

Yeast transformation was done with *S. cerevisiae* following Gietz protocol with small modifications (Gietz Lab Transformation Kit Manual, Molecular Research Reagents Inc.). *S. cerevisiae* was inoculated from yeast glycerol stock with sterile loop to 50 ml YPD medium in 250ml Erlenmeyer flask. The yeast was precultivated at 30 °C, 250 rpm overnight. 3 ml of the preculture was inoculated to 50 ml YPD medium and incubated at 30 °C, 250 rpm for around 3 h so that OD_{600} was around 1.

Next, the cells were collected (4 °C, 5000 rpm, 5 min), and resuspended to 20 ml sterile water by vortexing followed by centrifugation (4 °C, 5000 rpm, 5 min). The cells were resuspended again in 1 ml sterile water and centrifuged (4 °C, 10 000 rpm, 30 s). The supernatant was removed by pipetting. Then the cells were resuspended in 350 µl sterile water by pipetting up and down. 50 µl of the cell suspension was transferred into a 1.5 ml Eppendorf tube, centrifuged (10 000 rpm, 15 s) and the supernatant was removed.

Transformation mix containing 50 % PEG 4000 (240 µl), LiAc 1 M pH 7.5 (36 µl), ss-DNA (25 µl) and 100 ng plasmid DNA plus sterile water (50 µl) was added to the yeast cell

pellet. Then a heat shock was performed at 42 °C for 40 min in a heat block. After this, cells were centrifuged (5000 rpm, 1min) and resuspended in 1 ml sterile water. 100 µl of the cell suspension was plated to SCD-URA plate. The remaining 900 µl cell suspension was centrifuged (5000 rpm, 1min), the supernatant poured off and resuspended in the remaining supernatant around (100 µl). This cell suspension was also plated. Plates were incubated at 30 °C for 3-4 days.

6.6.2 Transformation of *E. coli* –Electroporation

3 µl of plasmid from Miniprep was added to 40 µl of electrocompetent cells. The mixture was transferred to electroporation cuvette and electroporation was performed at 1.5 kV. Then 1 ml SOC solution was added. 100 µl of the cell suspension was plated on LB plate with 100 µg/ml ampicillin. The remaining 900 µl cell suspension was centrifuged (5000 rpm, 1 min) and SOC solution was poured off. The cell pellet was resuspended in the remaining SOC solution (around 100 µl) and plated. The plates were incubated at 37 °C overnight.

6.6.3 Transformation of *A. niger* – Protoplast method

Protoplast method was used in the transformation of *A. niger* (Yelton *et al.*, 1984). First, protoplasts were prepared for the transformation. Sufficient amount (300-500 µl) of spores were inoculated to 125 YP-G medium in 250 Erlenmeyer flask and grown at 28 °C, 200 rpm for overnight. Then the germinated spores were filtered through a sterile Glass Microfibre Filter (55mm diameter, Whatman, GE Healthcare Life Science, USA) with vacuum and washed with 37 °C sterile water and cold KMC buffer (1 M KCl, 25 mM $CaCl_2$, 10 mM Tris-HCl, pH 5.8). Next, the mycelia were resuspended in 20 ml KMC buffer containing 200 mg caylase C4, incubated at 30 °C, 80 rpm for 3 h, and cooled on ice. The protoplasts were filtered through two layers of Miracloth (Calbiochem, Germany) prewet with cold KMC buffer. Filtration was done first without vacuum and towards the end with a soft vacuum. The protoplast solution was centrifuged (2 °C, 1500 x g, 4 min). The supernatant was poured off and the protoplasts washed with 20 ml cold KMC buffer by resuspending the pellet. Centrifugation was repeated and washed with 20 ml cold transformation buffer by resuspending the pellet. The used transformation buffer was STC solution (1.33 M sorbitol,10 mM Tris-HCl, 50 mM $CaCl_2$,

pH 8.0). The suspension was then centrifuged (2 °C, 1500 x g, 4 min), the supernatant removed and the pellet resuspended in STC solution 200 µl as the total volume.

Around 10 µg of each DNA altogether in 100 µl STC solution was used in the transformation. First 100 µl of the protoplasts were added to the DNA and then 100 µl of the transformation solution (25% PEG 6000, 50 mM $CaCl_2$, 10 mM Tris-HCl, pH 7.5) was added. The mixture was incubated on ice for 20 min. Then 2 ml transformation solution was added to the mixture and incubated at room temperature for 5 min. Next, 4 ml STC solution was added to stop the reaction. 7 ml melt top agar (50 °C) was added to the mixture and mixed by inverting the tube and poured to the selection plate (SCD-URA transformation plate or hygromycin (400 µg/ml) transformation plate) immediately. Plates were incubated at 28 °C for 3-4 days. The top agar consists of the same components as the transformation plate.

7 Results

7.1 Deletion of the genes potentially involved in the galactaric acid catabolism

In order to unravel the galactaric acid catabolic pathway based on transcriptomics data (not shown), eight genes, which were upregulated on galactaric acid and had a predicted function, were chosen for targeted deletion. Protein ID of these genes in Joint Genome Institute (JGI) are 39114, 1090836, 1117792, 1141260, 1121140, 1162477, 1146483 and 1170646 (Table 4).

Table 4 Proteins which are potentially involved in the galactaric acid catabolism in *A. niger*, and their predicted functions. Protein ID refers to the number of Joint Genome Institute.

Protein ID	Predicted function
39114	IPR000873: AMP-dependent synthetase and ligase
1090836	IPR001395: Aldo/keto reductase
1117792	IPR013149: Alcohol dehydrogenase, zinc-binding
1141260	IPR002198: Short-chain dehydrogenase/reductase SDR
1121140	IPR006076: FAD dependent oxidoreductase
1162477	IPR000623: Shikimate kinase, Gluconokinase
1146483	IPR013342: Mandelate racemase/muconate lactonizing enzyme, C-terminal
1170646	IPR006139: D-isomer specific 2-hydroxyacid dehydrogenase, catalytic region

The plasmids containing a deletion cassette were constructed to delete the eight potential genes involved in the galactaric acid catabolism from the Δ*pyrG* strain. A model of a plasmid containing a deletion cassette is shown in Figure 10. Each deletion cassette consists of specific flanking regions for the target gene and a *pyrG* gene as a selection marker for *A. niger*.

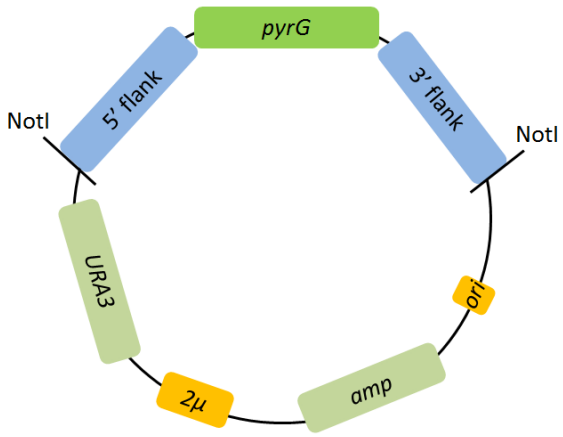

Figure 10 A model of a plasmid containing deletion cassette, which consists of the cassette and backbone plasmid p2974. The cassette consists of two 1.5 kb flanking regions amplified from *Aspergillus niger* genome and a *pyrG* gene as a selection marker for *A. niger*. The flanking regions are specific for the gene of interest. The size of the cassette is around 5 kb and designed to be released with NotI. The backbone p2974 (5.6 kb) have a bacteria and a yeast origin of replicon, o*ri* and *2μ* respectively, and genes *amp* (ampicillin resistance gene) and *ura3* are selection marker for *Escherichia coli* and *Saccharomyces cerevisiae* respectively.

The deletion of the eight potential genes related to the galactaric acid pathway was attempted from the Δ*pyrG*Δ*kusA* and Δ*pyrG* strains by homologous recombination. Transformants obtained from using the Δ*pyrG*Δ*kusA* strain are more likely to be correct compare to the transformant obtained from the Δ*pyrG* strain. This is due to the deletion of *kusA* gene involved in NHEJ which will result in the decrease of random integration of the deletion cassette. However, no transformants were obtained by transforming the Δ*pyrG*Δ*kusA* strain with the deletion cassettes. Therefore, only the Δ*pyrG* strain was manipulated.

The deletions were first performed by homologous recombination. The primary transformants were screened by colony PCR and purified by single spore isolation. Then the strains were cultivated on galactaric acid. Because clones of the Δ*39114* strain behaved differently when cultivated on galactaric acid, there was a worry about mixed population (mutants with correct deletion cassette integration and with random integration) also in other deletion mutants. Therefore, to confirm the correct integration of the deletion cassette and pureness of the culture, the 5' and 3' flanking regions of the cassette (Figure 11AB) and part of the ORF of the gene of interest (Figure 11C) were amplified. If both flanking regions and the ORF fragment were amplified the culture was assumed to not to be pure. If only the ORF fragment was amplified, the gene was not deleted. If only the cassette fragments were amplified, the culture was pure. The Δ*39114* strain was only tested by amplifying the 5' flanking region and the ORF, but the 3' flanking region was not tested. However, the Δ*39114* strain was confirmed to be pure, because a clear phenotype could be detected when it was cultivated with galactaric acid. The pureness of the strain was later also confirmed by colony PCR.

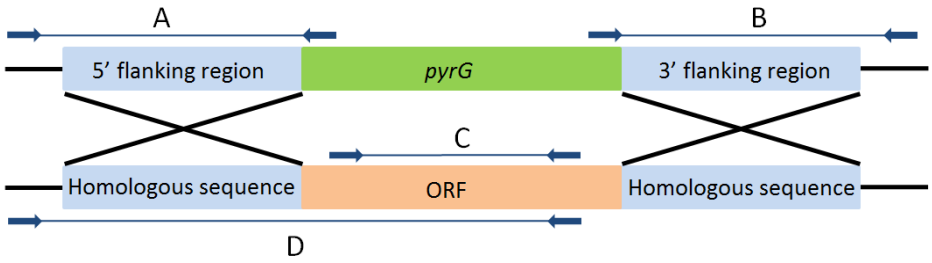

Figure 11 Fragments amplified in colony PCR for screening and purifying the deletion strains. Arrows represent the primers used in PCR. Size of the fragments are A) around 1.5 kb, B) around 1.7 kb, C) 700-770 bp and D) 2.3-2.8 kb depending on the gene.

Generally, it was difficult to obtain pure strains when the deletion was performed by homologous recombination. All in all, identifying correct colonies and going through singles spores to purify the stains was done 3-7 times depending on the deletion. In most cases, 5' and 3' flanking region and ORF fragments were amplified (Figure 12 no2), indicating that in the mutant the gene was not deleted, but the cassette was still integrated to the correct place. In some cases, only 5' flanking region and ORF was amplified (Figure 12 no1). In these cases, maybe the gene was only partially deleted. Some deletion mutants were also lost during screening and going through single spores. Only Δ*39114* was purified with single spore isolation when the deletion was performed by homologous recombination.

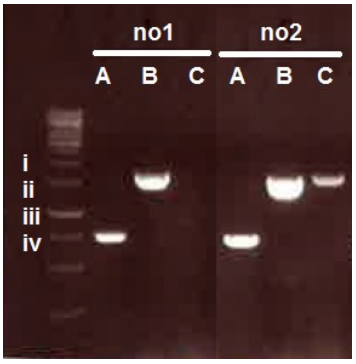

Figure 12 A gel picture of colony PCR when deleting genes by homologous recombination. Fragments size in the ladder are i = 2 kb, ii = 1.5 kb, iii = 1 kb and iv = 750 bp. The amplified fragments in colony PCR were A = ORF (700-770 bp), B = 5' flanking region (1.5 kb) and C = 3' flanking region (1.7 kb).

After the several attempts to delete the genes by homologous recombination and to purify the mutant strains, deletion with the CRISPR/Cas9 method using deletion cassettes as a donor DNA was performed. 8 primary transformants per deletion were screened by colony PCR (and plated to SCD-URA), but again ORF and both flanking regions were amplified. There were three different cases: only ORF, ORF+5' flanking region and ORF+5'+3' flanking regions (Figure 13 no1, no3, no2). Thus, colony PCR was performed with few colonies in which all three fragments were amplified, using primers binding outside of the 5' flanking region and inside the ORF (Figure 11D). If nothing could be amplified, the gene should not be in its original place and, thus, the gene should have been deleted. In some colonies this fragment was amplified and in some

not. After the primary transformants had gone through single spore isolation once, most of the clones did not have the ORF anymore and only the 5' and 3' flanking region of the deletion cassette were amplified (Figure 13 no4). After going through the single spore isolation for the second time, the ORF fragment was not amplified and 3' flanking region was also not amplified in some clones. Nevertheless, in at least half of the screened clones only the flanking regions were amplified indicating that the deletion was correct. In total, the transformants were replated six times, alternating between SCD-URA and PD plates in order to screen and purify the strain by single spore isolation.

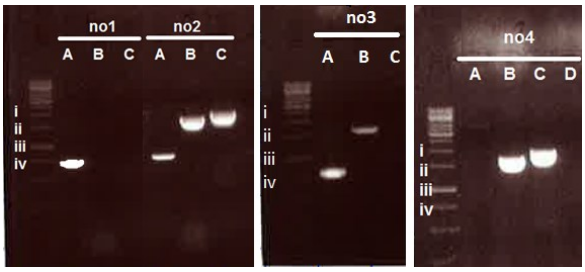

Figure 13 A gel picture of colony PCR when deleting genes using the CRISPR/Cas9 method. Fragments size in the ladder are i = 2 kb, ii = 1.5 kb, iii = 1 kb and iv = 750 bp. The amplified fragments in colony PCR were A = the ORF (700-770 bp), B = 5' flanking region (1.5 kb), C = 3' flanking region (1.7 kb) and D = 5' flanking region and part of the ORF (2.3-2.8 kb depending on the gene).

The use of CRISPR/Cas9 for gene deletion increased the deletion efficiency significantly. Comparison of deleting genes *1090836*, *1117792*, *1141260* and *1121140* by homologous recombination with deletion cassette and by the CRISPR/Cas9 method using two gRNAs and deletion cassette are shown in Table 5. With the CRISPR/Cas9 method, the efficiency was increased 8-fold when deleting the *1090836* gene and 2-fold when deleting the *1117792* gene. The homologous recombination efficiency was very low when deleting the genes *1141260* and *1121140*. None of the 30 screened Δ*1141260* primary transformants had the gene deleted. In the case of the Δ*1121140* primary transformants, only one primary transformant of the 60 screened transformants was correct. In fact, performing the deletion of the gene *1121140* again resulted in 0 correct transformants out of the 60 screened transformants. On the other hand, when the deletion was done with the CRISPR/Cas9 method, all 8 screened Δ*1141260* transformants were correct and 3 out of 8 screened Δ*1121140* transformants were correct.

Table 5 Comparison of deleting four different genes with homologous recombination and the CRISPR/Cas9 mthod. Amplifying flanking regions of the deletion cassette by colony PCR was used to verify the correct transformants. Efficiency is percentage of correct transformants of total screened colonies.

	Homologous recombination			CRISPR/Cas9		
Deletion	Total	Correct	Efficiency (%)	Total	Correct	Efficiency (%)
1090836	30	1	3.3	40	11	27.5
1117792	30	13	43.3	8	8	100
1141260	30	0	0.0	8	8	100
1121140	60	1	1.7	8	3	37.5

7.2 New *A. niger* phenotype of lacking galactaric acid catabolism

In order to screen new phenotypes, after the successful deletion of the genes and strain purification, two clones of each mutant stain, Δ*39114*, Δ*1090836*, Δ*1117792*, Δ*1141260*, Δ*1121140*, Δ*1162477*, Δ*1146483* and Δ*1170646,* were cultivated with two replicates for 5 days in three different media, MM-McA, MM-XMcA and YP-McA, containing 10 g/l of galactaric acid. The consumption of galactaric acid was analysed by HPLC. Mutants which did not consume galactaric acid were concluded to have a phenotype that was not described previously.

The consumption of galactaric acid by the wild-type and the mutant strains was poor in MM-McA medium (Figure 14). Deletion of the genes *39114*, *1090836*, and *1121140* resulted in strains, which did not consume galactaric acid when cultivated in MM-XMcA medium (Figure 15). In YP-McA medium, the consumption between the wild-type strain and the Δ*39114* strain was clearly different (Figure 16). However, the Δ*1090836* and Δ*1121140* strains consumed galactaric acid in YP-McA, but the consumption was less compared to the wild-type strain. The wild-type strain consumed almost all galactaric acid in MM-XMcA and YP-McA media. Other deletion mutants, Δ*1117792*, Δ*1141260*, Δ*1162477*, Δ*1146483* and Δ*1170646* consumed galactaric acid and behaved more or less similar to the wild-type strain in all culture media. The increase in the galactaric acid concentration in the cultures was due to the evaporation of the medium during cultivation. The Δ*39114* strain was also cultivated with glucaric acid and it did not consume the acid, whereas the wild-type strain consumed all (Figure 17). The Δ*1162477* was also grown on gluconic acid plates. It was assumed that the mutant might not grow, because the gene has homology to a gluconokinase and this activity was assumed to be

essential for gluconic acid catabolism. However, the growth of the Δ*1162477* strain looked the same as the wild-type strain.

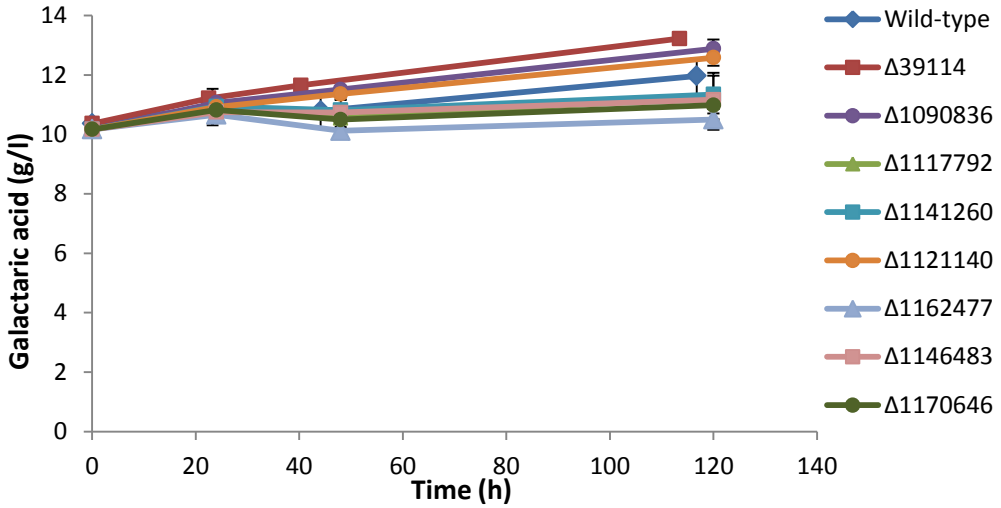

Figure 14 Galactaric acid consumption of the wild-type strain ATCC 1015 and the Δ*39114*, Δ*1090836*, Δ*1117792*, Δ*1141260*, Δ*1121140*, Δ*1162477*, Δ*1146483* and Δ*1170646* strains during 5-day cultivation. The stains were cultivated in minimal medium with 10 g/l galactaric acid (MM-McA). Average of two clones and their two replicates are presented in the figure. Error bars represent ±standard deviation. The wild-type strain also has error bar of the time, because half of the samples were taken at significantly different time.

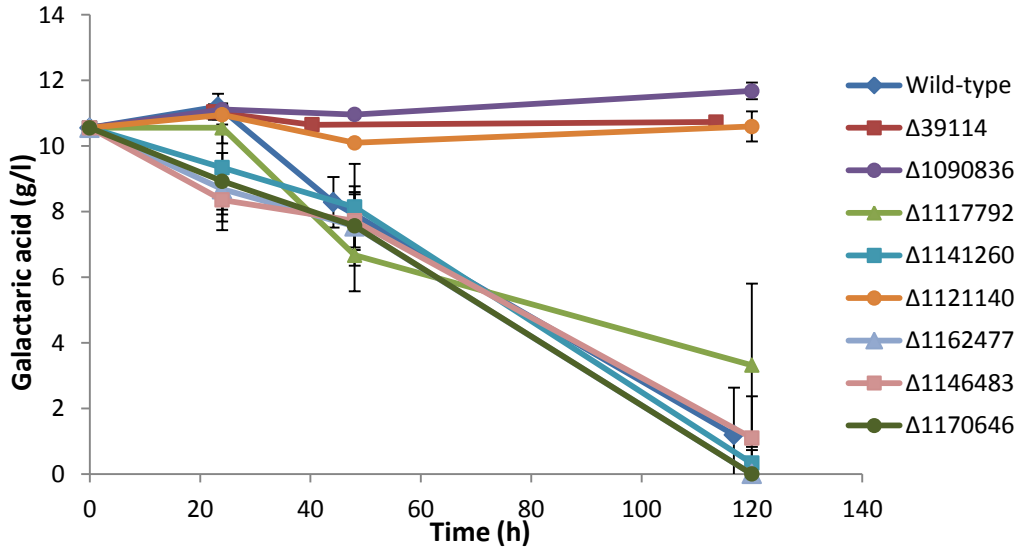

Figure 15 Galactaric acid consumption of the wild-type strain ATCC 1015 and the Δ*39114*, Δ*1090836*, Δ*1117792*, Δ*1141260*, Δ*1121140*, Δ*1162477*, Δ*1146483* and Δ*1170646* strains during 5-day cultivation. The stains were cultivated in minimal medium with 0.5 % xylose and 10 g/l galactaric acid (MM-XMcA). Average of two clones and their two replicates are presented in the figure. Error bars represent ±standard deviation. The wild-type strain also has error bar of the time, because half of the samples were taken at significantly different time.

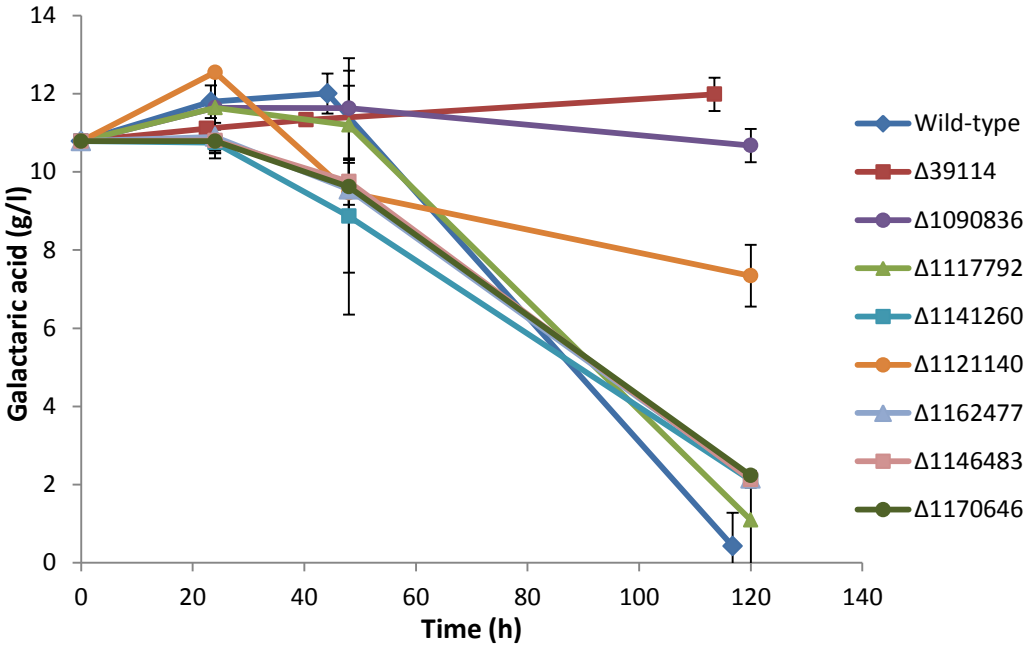

Figure 16 Galactaric acid consumption of wild-type strain ATCC 1015 and the Δ*39114*, Δ*1090836*, Δ*1117792*, Δ*1141260*, Δ*1121140*, Δ*1162477*, Δ*1146483* and Δ*1170646* strains during 5-day cultivation. The stains were cultivated in YP medium with 10 g/l galactaric acid (YP-McA). Average of two clones and their two replicates are presented in the figure. Error bars represent ±standard deviation. The wild-type strain also has error bar of the time, because half of the samples were taken at significantly different time.

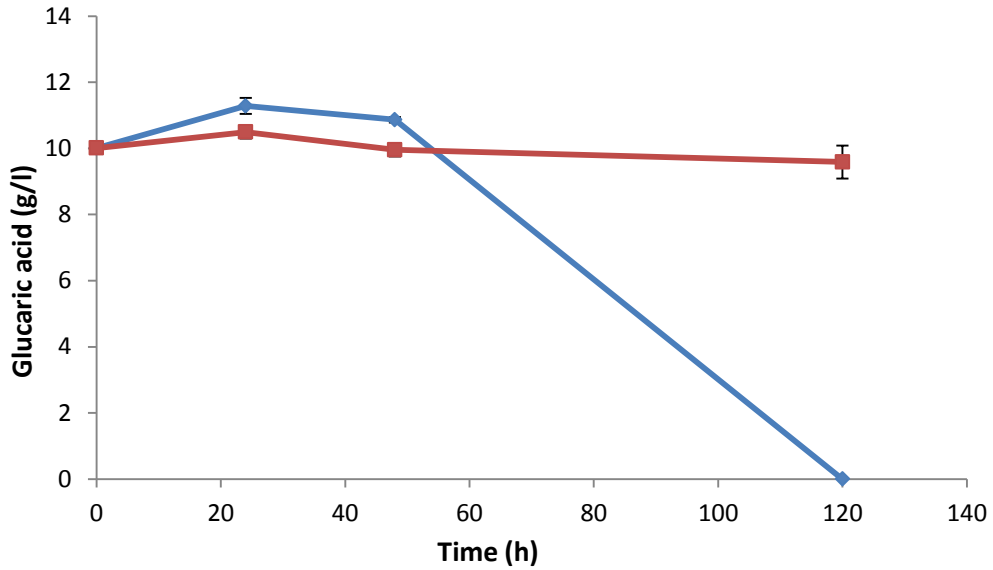

Figure 17 Glucaric acid consumption of the wild-type strain ATCC 1015 (◆) and the Δ*39114* strain (■) during 5-day cultivation in YP medium with 10 g/l glucaric acid (YP-GlcA). Average of two replicates of wild-type and average of two clones and their two replicates of the Δ*39114* strain are presented in the figure. Error bars represent ±standard deviation. If the error bars are not visible, they are smaller than the symbols.

The colour change of the cultures was observed during the galactaric cultivation (Table 6). The colour of the wild-type strain cultures were yellow in all media. After 5 days, media of MM-McA cultures of the strains Δ*1090836*, Δ*1141260*, Δ*1121140*, Δ*1146483* and Δ*1170646* and one clone of the Δ*1117792* strain, had turned reddish, whereas the strains Δ*39114* and Δ*1162477* and another clone of Δ*1117792* were yellow. In the case of MM-XMcA cultures, the Δ*1090836* culture turned to orange, the Δ*1121140* culture to orangish yellow and the rest to yellow. Media of YP-McA cultures remained yellow except Δ*39114*, Δ*1090836* and Δ*1121140* turned to dark orange or orange. The reason for the colour change is unknown. Probably the deletion had caused the formation and maybe also accumulation of some colour producing compounds.

Table 6 The colour of galactaric acid cultures of different mutant strains after 5-day cultivation.

Strain/Media	**MM-McA**	**MM-XMcA**	**YP-McA**
Wild-type	Yellow	Yellow	Yellow
Δ*39114*	Yellow	Yellow	Orange
Δ*1090836*	Reddish	Orange	Dark orange
Δ*1117792*	Reddish/Yellow	Yellow	Yellow
Δ*1141260*	Reddish	Yellow	Yellow
Δ*1121140*	Reddish	Orangish yellow	Orange
Δ*1162477*	Yellow	Yellow	Yellow
Δ*1146483*	Reddish	Yellow	Yellow
Δ*1170646*	Reddish	Yellow	Yellow

In order to express genes *39114*, *1090836* and *1121140*, which deletion resulted in the strains not able to utilize galactaric acid, in *S. cerevisiae* and measure the activity of the enzymes coded by these genes, the ORF of the genes *39114*, *1090836* and *1121140* were amplified from the cDNA and and ligated with the linearized yeast expression vector p2159. Then the whole ORF was sequenced to see whether intron prediction of JGI was correct. Intron prediction of JGI for the genes *39114* and *1090836* differed from the sequencing results. The gene *1121140* does not have introns as predicted in JGI. The sequence of the genes with introns are shown in Appendix 2. The ORF of the genes

39114, *1090836* and *1121140* was expressed in *S. cerevisiae* CEN.PK2-1D. Enzyme activity of the enzyme 39114 was tested. The activity of the 39114 enzyme was not detected when galactaric or glucaric acid were used as the substrate with cofactors: NADP, NADPH, NAD and NADH.

7.3 Construction and cultivation of galactaric acid producing strains

7.3.1 Construction of the Δ*39114*Δ*gaaA/C*::*udh* and Δ*39114*Δ*gaaA/C udh* strains

In order to construct an *A. niger* strain which can produce galactaric acid, D-galacturonic acid and galactaric acid catabolic pathways must be blocked and D-galacturonate dehydrogenase for galactaric acid production must be expressed (Figure 18). Two different approaches for the deletion of the genes *gaaA* and *gaaC* (*gaaA/C*) coding for D-galacturonate reductase and 2-keto-3-deoxy-L-galactonate aldolase respectively in the D-galacturonic acid catabolic pathway from the strain Δ*39114* were developed. First, the deletion was attempted by integrating the *udh* gene coding for D-galacturonate dehydrogenase derived from *P. syringae* to the *gaaA*/*C* locus by homologous recombination using an *udh* integration cassette (resulst in Δ39114Δ*gaaA/C*::*udh* strain). Secondly, the CRISPR/Cas9 method was utilized to delete *gaaA/*C (results in Δ39114Δ*gaaA/C* strain) and then randomly integrate the *udh* gene to the genome simultaneously or later (results in Δ39114Δ*gaaA/C udh* strain). In the genome, the genes *gaaA* and *gaaC* are located next to each other separated only by a bidirectional promoter. Therefore, simultaneous deletion of the genes would be possible.

In order to delete the *gaaA/C* from the Δ*39114* strain and express the *udh*, construction of a plasmids containing the *udh* integration cassette was attempted. The fragments of the plasmid (two flanking regions, *udh* and backbone plasmid) were amplified by PCR and attempted to assemble using homologous recombination in *S. cerevisiae*, TOPO Cloning or Gibson Assembly. In TOPO Cloning, the amplification of integration cassette as one fragment occurred because of the 40 bp overhangs of the templates capable of annealing to each other The integration cassette could not be obtained with any of the tried methods. When assembling integration cassette using yeast homologous recombination and Gibson Assembly, only the 5’ flanking region was amplified by colony PCR. When using TOPO Cloning both flanking region fragments could not be amplified.

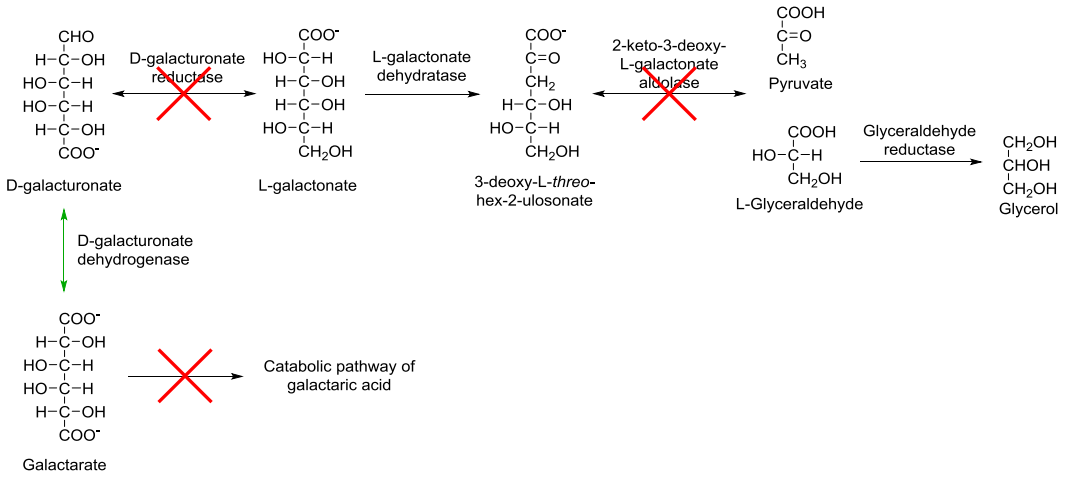

Figure 18 Catabolism of D-galacturonic acid in *A. niger* can be blocked by deleting D-galacturonate and 2-keto-3-deoxy-L-galactonate aldolase of the catabolic pathway. Expressing a D-galacturonate dehydrogenase would enable the production of galactaric acid in *A. niger*, if the galactaric acid catabolic pathway is also blocked.

Because the *udh* integration cassette could not be obtained with yeast homologous recombination, TOPO Cloning nor Gibson Assembly, the CRISPR/Cas9 method was applied for simultaneous deletion of the *gaaA/C* gene and the integration of the *udh* gene in the Δ*39114* strain and inactivation of the *gaaA/C* gene in the Δ*39114* strain. When using the CRISPR/Cas9 method in gene deletion, 1 µg of plasmid pFC-332 was enough for transformation. The transformation was also done with 10 µg, but it led to the overgrowth of the transformants. The plasmid pFC-332 is an autonomously replicating vector harboring AMA1 (autonomous maintenance in *Aspergillus*) sequence and a *hph* (hygromycing resistance) gene as a selection marker. Under the selective conditions, the plasmid pFC-332 is maintained in the cell, but readily lost without selection. Least primary transformants were obtained in the Δ*39114ΔgaaA/C::udh* strain, more with the Δ*39114ΔgaaA/C* strain and most with the control strain. The Δ*39114ΔgaaA/C* transformants were screened by colony PCR. The PCR fragment size of colonies, in which the *gaaA/C* gene is deleted, should be around 200 bp assuming there were no indels. Integration of the *udh* gene to the *gaaA/C* locus would results in a 2691 bp fragment and if the *gaaA/C* gene was not deleted, the fragment size would be 3406 bp. In 9 out of 26 screened colonies, the *gaaA/C* (3.4 kb) was not amplified, but the 200 bp fragment could also not be observed. These results does not confirm the deletion of the *gaaA/C* gene. However, the 200 bp fragment is quite small to be seen on the gel.

Therefore, these colonies were also screened by cultivating on SCGalA-URA plates. The deletion of the *gaaA/C* gene should result in the strain not being able to grown on galacturonic acid. However, the clones were still able to grow on galacturonic acid suggesting that the *gaaA/C* gene was not deleted. A possible explanation to the colony PCR and growth results could be that the *gaaA*/*C* gene was deleted, but this deleted cluster integrated back to the genome randomly to a different place. The Δ*39114*Δ*gaaA/C::udh* transformants were also screened by colony PCR. Totally, 52 colonies were screened but the *udh* gene was not amplified. The results showed again that the *gaaA/C* gene was still in the genome or that it was possibly deleted from its original locus. As a conclusion, with this approach the *gaaA/C* gene was not deleted nor the *udh* gene integrated.

7.3.2 Construction of the Δ*gaaA udh* Δ*39114* strain

Since the strains Δ*39114*Δ*gaaA/C::udh* and Δ*39114*Δ*gaaA/C udh* could not be constructed due to unsuccessful construction of the *udh* integration cassette and deletion of the *gaaA/C* gene, a new approach was applied. In this approach, the *39114* gene was deleted from the Δ*gaaA udh* stain using the CRISPR/Cas9 method. The deletion was performed in two different ways, using two gRNAs and one gRNA with the deletion cassette as a donor DNA. Least primary transformants were obtained in the transformation with the Δ*gaaA udh* strain with two gRNAs, then the control transformation and the Δ*gaaA udh* strain with one gRNA and deletion cassette. 24 primary transformants of the construction of the Δ*gaaA udh* Δ*39114* strain without the cassette were screened, but none of these transformants had shorterned ORF, as would be expected if the gene was inactivated with two DSBs. 2 out of the 24 screened primary transformants of the construction of the *ΔgaaA udh Δ39114* strain with deletion cassette were correct. In these transformants, the ORF and the 5' and the 3' flanking regions were amplified in colony PCR. During the purification of the one correct transformant, only the 5' and 3' flanking regions were amplified and the ORF was not, indicating the successful deletion of the *39114* gene and that the culture is pure. In the purification of another correct transformants after going through single spores once, only the ORF was amplified or in 2 out of the8 screened colonies the ORF and the 3' flanking region was amplified.

7.3.3 Cultivation of the Δ*gaaA udh* Δ*39114* strain

Two clones of the strain Δ*gaaA udh* Δ*39114* and the parent strain Δ*gaaA udh* were grown on SCGalA-URA plate (Figure 19). After incubating the galacturonic acid plate at 28 °C for 3 days, the parent strain Δ*gaaA udh* had formed a white colony and sporulated and the Δ*gaaA udh* Δ*39114* strain had formed a yellow colony without any spores. The Δ*gaaA udh* Δ*39114* strain grew slower than Δ*gaaA udh* strain. The size of the Δ*gaaA udh* Δ*39114* colony is also smaller and much thinner than the Δ*gaaA udh* colony.

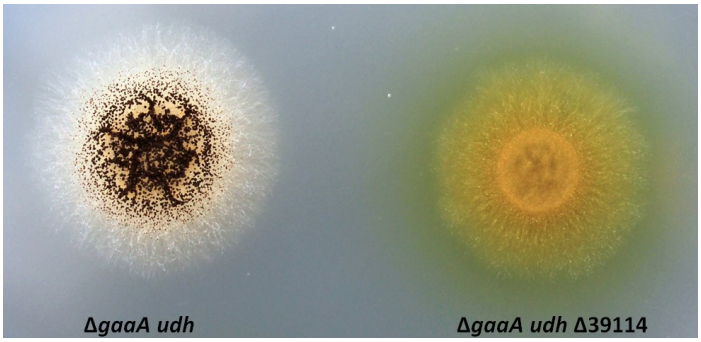

Figure 19 Cultivation of the parent strain Δ*gaaA udh* and the mutant strain Δ*gaaA udh* Δ*39114* on SCGalA-URA agar plate, 28 °C, 3 days. The Δ*gaaA udh* strain had formed a white colony which had started to sporulate, whereas the Δ*gaaA udh* Δ*39114* strain had formed a smaller and thinner yellow colony without spores.

The strains Δ*gaaA udh* Δ*39114* and Δ*gaaA udh* were also cultivated in submerged cultivations in different scales in different media. Initial concentration of galacturonic acid was 20 g/l in all cultivations. In 4 ml-scale cultivation, the strains Δ*gaaA udh* Δ*39114* and Δ*gaaA udh* were cultivated in MM-GalA medium (Figure 20) and MM-XGalA medium (Figure 21) with initial pH of 5. The Δ*gaaA udh* culture had turned to yellow in both medium after 5 days and the Δ*gaaA udh* Δ*39114* culture had turned light yellow. In MM-GalA cultivations, the Δ*gaaA udh* strain consumed galacturonic acid, but did not produced galactaric acid. The Δ*gaaA udh* Δ*39114* strain did not consume much galacturonic acid, but was able to produce 0.8 ± 0.03 g/l of galactaric acid. The increase of galacturonic acid is due to the evaporation of the medium during cultivation. The evaporation might have also affected galactaric acid concentration. In MM-XGalA cultivations, similar patterns of galactaric acid consumption as in the MM-GalA cultivations can be seen in both strains. However, both strains were able to produce galactaric acid. After 48 h the Δ*gaaA udh* strain produced 0.5 ± 0.02 g/l galactaric acid,

but it was all consumed after 120 h. Whereas the Δ*gaaA udh* Δ*39114* strain produced 1.1 ± 0.04 g/l of galactaric acid.

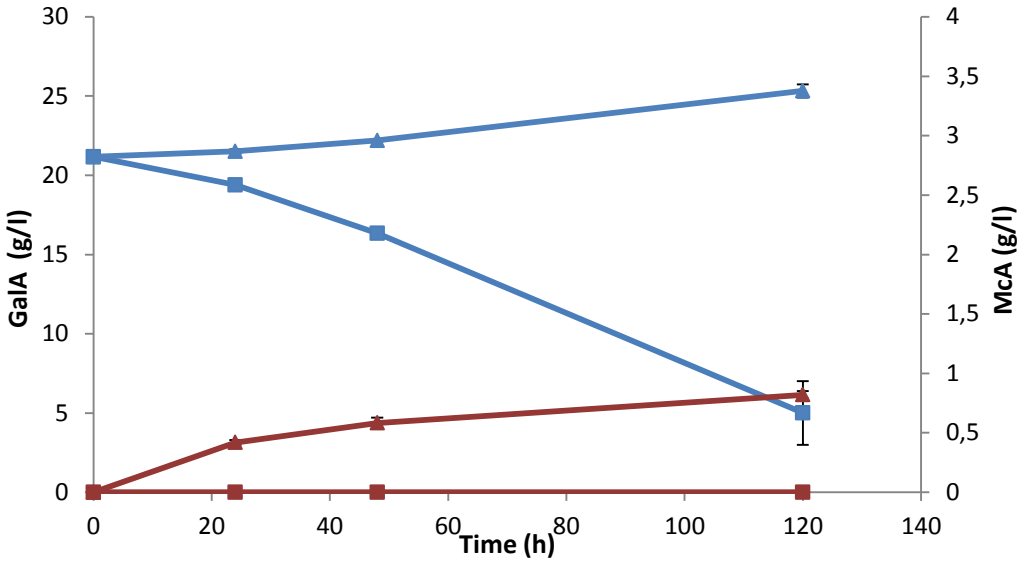

Figure 20 Consumption of galacturonic acid (GalA, blue symbols) and formation of galactaric acid (McA, red symbols) during 5-day cultivation in 4 ml MM-GalA medium (pH 5) by the strains Δ*gaaA udh* (■) and Δ*gaaA udh* Δ*39114* (▲). Average of two replicates of the Δ*gaaA udh* and average of two clones and their two replicates of the Δ*gaaA udh* Δ*39114* are presented in the figure. Error bars represent ±standard deviation. If the error bars are not visible, they are smaller than the symbols.

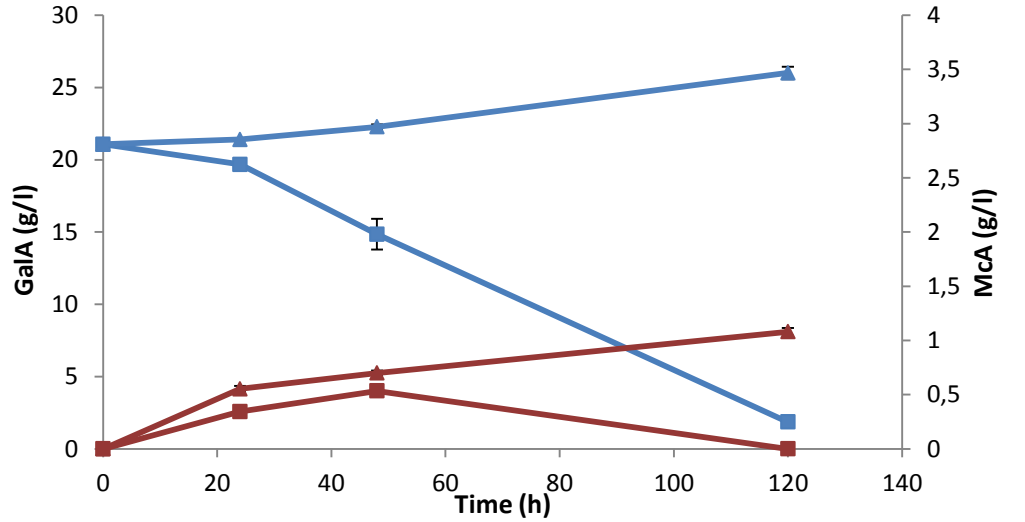

Figure 21 Consumption of galacturonic acid (GalA, blue symbols) and formation of galactaric acid (McA, red symbols) during 5-day cultivation in 4 ml MM-XGalA medium (pH 5) by the strains Δ*gaaA udh* (■) and Δ*gaaA udh* Δ*39114* (▲). Average of two replicates of the Δ*gaaA udh* and average of two clones and their two replicates of the Δ*gaaA udh* Δ*39114* are presented in the figure. Error bars represent ±standard deviation. If the error bars are not visible, they are smaller than the symbols.

In 50 ml-scale submerged cultivation, the strains Δ*gaaA udh* Δ*39114*, Δ*gaaA udh* and wild-type were cultivated in the MM-GalA and the YP-GalA media with initial pH 5 or pH 3 and in the MM-XGalA medium with initial pH 5 for 5 days. All MM-GalA cultures with initial pH 5 turned to light yellow after 24 h, and after 120 h the wild-type and the Δ*gaaA udh* cultures were brown whereas the Δ*gaaA udh* Δ*39114* cultures were still yellow. With the initial pH of 3, all MM-GalA cultures were clear after 24 h, and after 120 h the wild-type cultures was brownish orange, the Δ*gaaA udh* cultures had brown cells and yellow supernatant and the Δ*gaaA udh* Δ*39114* cultures was light yellow. The MM-XGalA cultures of the wild-type and the Δ*gaaA udh* strains were brown and the Δ*gaaA udh* Δ*39114* strain were yellow brown. The colour of the YP-GalA cultures of the wild-type, the Δ*gaaA udh* and the Δ*gaaA udh* Δ*39114* strains after 120 h with initial pH 5 were yellow brown, yellow brown and dark brown, and with initial pH 3 the cultures were dark brown, brown and brown yellow, respectively. In general, the size of the cell was the smallest in the Δ*gaaA udh* Δ*39114* cultures, and the wild-type strain had produced most biomass.

The consumption of galacturonic acid and the production of galactaric acid during the cultivations are shown in Figure 22 – Figure 26. In general, the wild-type strain consumed all galacturonic acid after 48 h and the Δ*gaaA udh* strain consumed all or almost all after 48, 72 or 120 h. The Δ*gaaA udh* Δ*39114* strain consumed very little galacturonic acid in the MM-GalA media, but a clear consumption could be seen in the YP-GalA media. The wild-type strain did not produce galactaric acid in any medium. The Δ*gaaA udh* strain did not produce galactaric acid in the MM-GalA with initial pH 5, however with initial pH 3 it produced 0.3 ± 0.009 g/l galactaric acid after 48 h. Nevertheless, all produced galactaric acid was consumed after 120 h. The Δ*gaaA udh* Δ*39114* strain was able to produce 1.5 ± 0.03 g/l galactaric acid after 120 h in the MM-GalA medium with initial pH 5 and 0.9 ± 0.02 g/l in the MM-GalA medium with initial pH 3. In the MM-XGalA medium with initial pH 5, the Δ*gaaA udh* strain produced 0.4 ± 0.009 g/l galactaric acid after 24 h, but they were all consumed after 48 h. The Δ*gaaA udh* Δ*39114* strain produced 1.5 ± 0.5 g/l galactaric acid after 120 h which is almost same as in the MM-GalA medium. The production rate was, however, higher in the MM-XGalA medium compared to in the MM-GalA medium. In the YP-GalA medium compared to mineral media, galactaric acid production was improved. In the YP-GalA medium with initial pH of 5, the Δ*gaaA udh* strain produced 1.2 ± 0.02 g/l and the

Δ*gaaA udh* Δ*39114* strain 4.3 ± 0.2 g/l galactaric acid after 120 h. With initial pH of 3 the production was 1.0 ± 0.02 g/l and 3.4 ± 0.1 g/l, respectively.

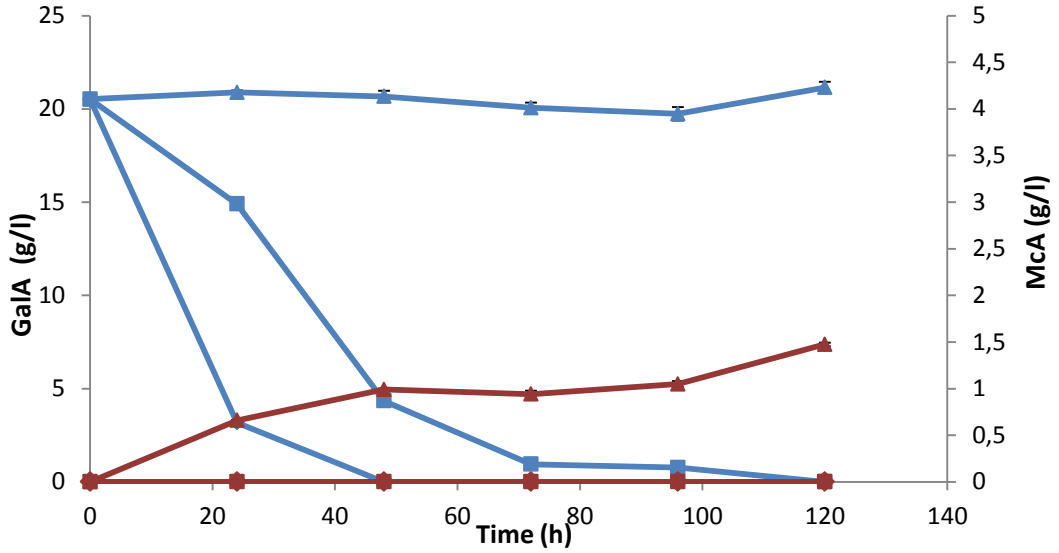

Figure 22 Consumption of galacturonic acid (GalA, blue symbols) and formation of galactaric acid (McA, red symbols) during 5-day cultivation in 50 ml MM-GalA medium, pH 5 by strains wild-type (◆), Δ*gaaA udh* (■)and Δ*gaaA ud*h Δ*39114* (▲). Average of three replicates is presented in the figure. Error bars represent ±standard deviation. If the error bars are not visible, they are smaller than the symbols.

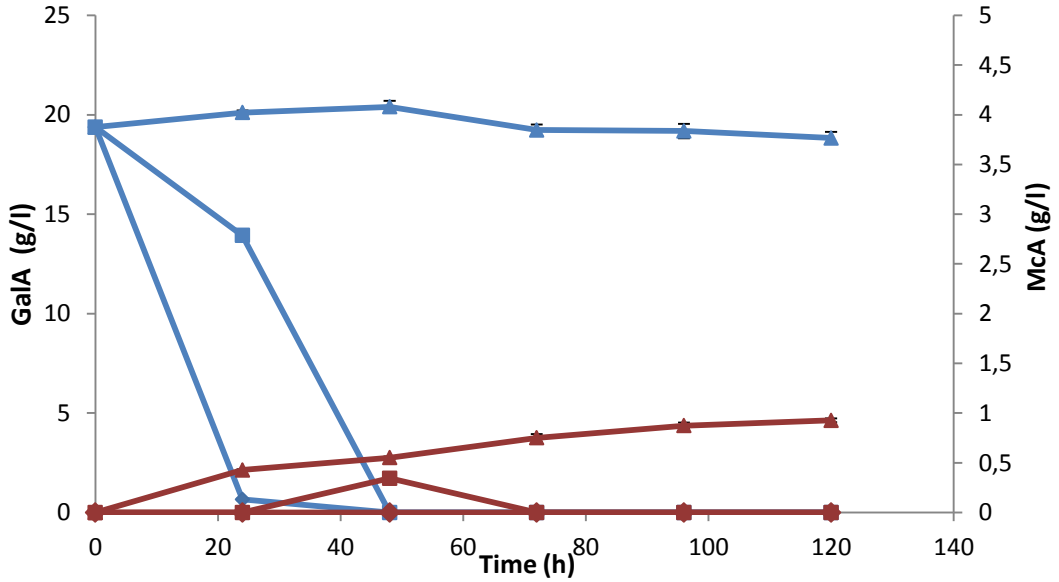

Figure 23 Consumption of galacturonic acid (GalA, blue symbols) and formation of galactaric acid (McA, red symbols) during 5-day cultivation in 50 ml MM-GalA medium, pH 3 by strains wild-type (◆), Δ*gaaA udh* (■)and Δ*gaaA ud*h Δ*39114* (▲). Average of three replicates is presented in the figure. Error bars represent ±standard deviation. If the error bars are not visible, they are smaller than the symbols.

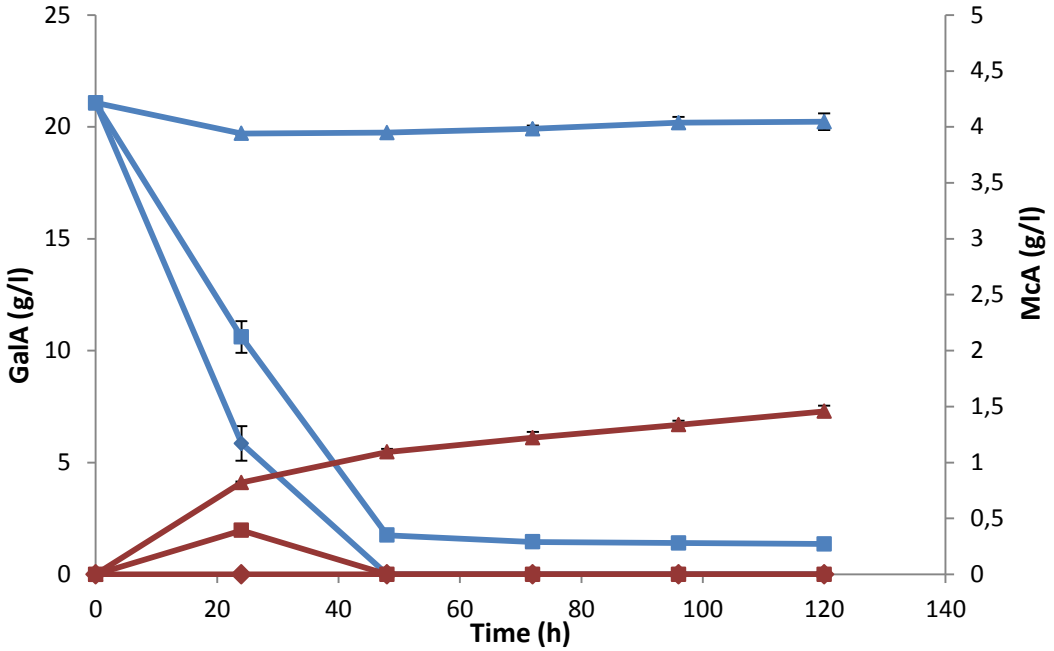

Figure 24 Consumption of galacturonic acid (GalA, blue symbols) and formation of galactaric acid (McA, red symbols) during 5-day cultivation in 50 ml MM-XGalA medium, pH 5 by strains wild-type (◆), Δ*gaaA udh* (■)and Δ*gaaA ud*h Δ*39114* (▲). Average of three replicates is presented in the figure. Error bars represent ±standard deviation. If the error bars are not visible, they are smaller than the symbols.

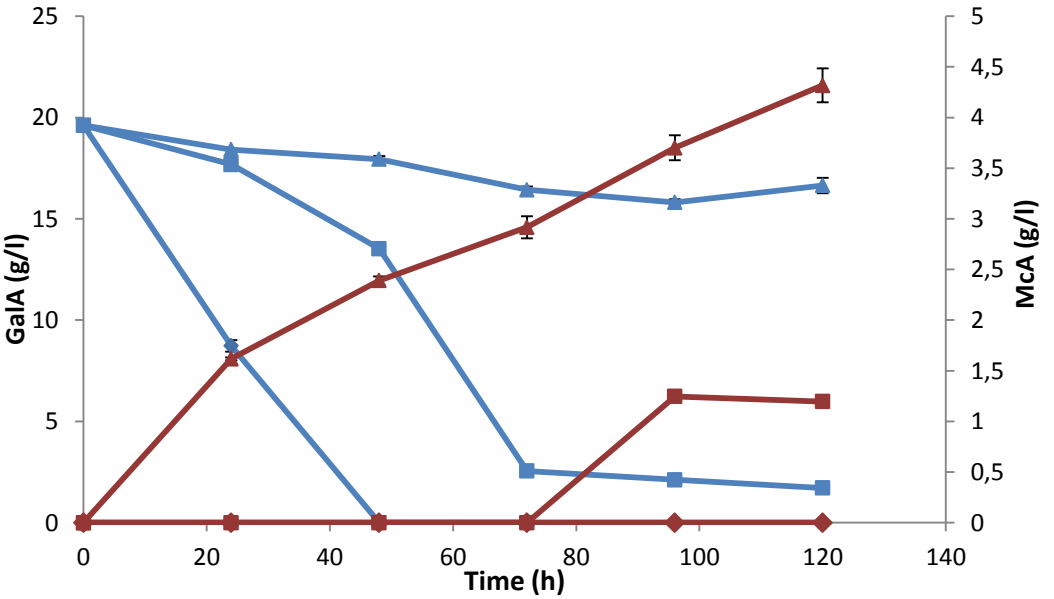

Figure 25 Consumption of galacturonic acid (GalA, blue symbols) and formation of galactaric acid (McA, red symbols) during 5-day cultivation in 50 ml YP-GalA medium, pH 5 by strains wild-type (◆), Δ*gaaA udh* (■)and Δ*gaaA ud*h Δ*39114* (▲). Average of three replicates is presented in the figure. Error bars represent ±standard deviation. If the error bars are not visible, they are smaller than the symbols.

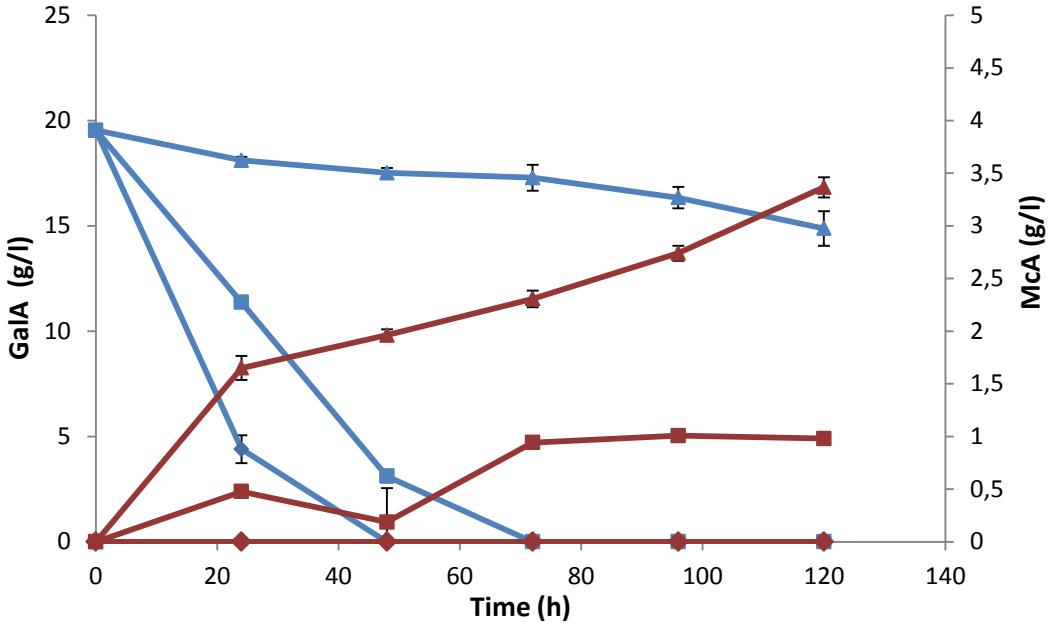

Figure 26 Consumption of galacturonic acid (GalA, blue symbols) and formation of galactaric acid (McA, red symbols) during 5-day cultivation in 50 ml YP-GalA medium, pH 3 by strains wild-type (◆), Δ*gaaA udh* (■)and Δ*gaaA ud*h Δ*39114* (▲). Average of three replicates is presented in the figure. Error bars represent ±standard deviation. If the error bars are not visible, they are smaller than the symbols.

The Δ*gaaA udh* Δ*39114* strain was also fermented in SSF and SmF using CPW as the substrate. The wild-type strain and the Δ*gaaA udh* strain were also fermented in SmF using CPW as the substrate. After keeping 4.0438 g of dried CPW at 100 °C, it mass had reduced to 3.7858 g. Therefore, 1 g CPW represents 0.936 g on a DM due to residual water. In the SSF using CPW as a substrate, the Δ*gaaA udh* Δ*39114* strain was able to produce 16.7 ± 0.9 mg galactaric acid from 2 g (1.87 g DM) CPW (8.9 ± 0.5 mg galactaric acid/g DM CPW) in 137 h. When nitrogen was also supplemented, the production increased to 28.9 ± 1.1 mg galactaric acid from 2 g CPW (15.4 ± 0.6 mg galactaric acid/g DM CPW). In the SmF, the peel waste was degraded after 24 h and no particles were visible. The wild-type strain and the Δ*gaaA udh* strain did not produce galactaric acid, but the Δ*gaaA udh* Δ*39114* strain had produced 3.1 ± 0.1 g/l galactaric acid from 40 g/l (37.4 g/l DM) CPW (82.3 ± 3.9 mg galactaric acid/g DM CPW) after 120 h (Figure 27). Galacturonic acid was probably released to the liquid faster in the fermentation with the wild-type strain and the Δ*gaaA udh* strain than with the Δ*gaaA udh* Δ*39114* strain. The wild-type strain and the Δ*gaaA udh* strain also

consumed the galacturonic acid faster. In the fermentation with the Δ*gaaA udh* Δ*39114* strain after 120 h, 8.4 ± 0.06 g/l galacturonic acid was released.

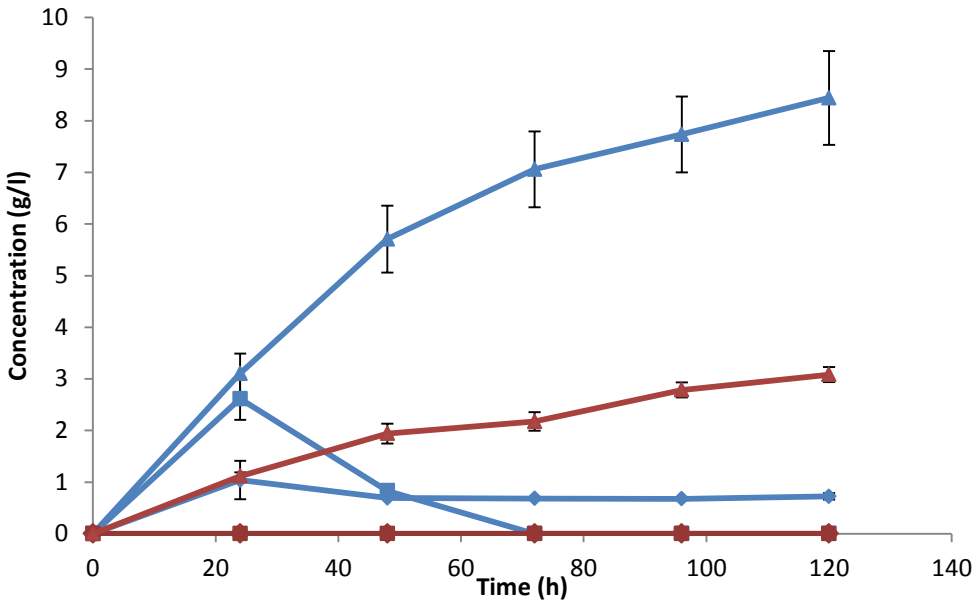

Figure 27 Submerged fermentation of the strains wild-type (◆), Δ*gaaA udh* (■) and Δ*gaaA udh* Δ*39114* (▲) using citrus processing waste as the substrate. The figure shows the release of galacturonic acid (GalA,blue) and the production of galactaricacid (McA,red). The wild-type strain and the Δ*gaaA udh* did not produce galactaric acid.

8 Discussion

8.1 Strategies for gene deletion in *A. niger*

The deletion of potential genes involved in galactaric acid pathway was attempted using two different strains, the Δ*pyrG* and the Δ*pyrG*Δ*kusA* strains, and two methods, homologous recombination and CRISPR/Cas9. The *kusA* deletion should have facilitated the homologous recombination, however, no transformants were obtained by transforming the Δ*pyrG*Δ*kusA* strain with the deletion cassettes. It has been reported that deleting the *kusA* gene might result in an unstable strain due to the weakened ability to repair DSBs. SSBs, on the other hand, can be repaired by KusA independent repair mechanisms. The Δ*kusA* strain is more sensitive towards UV and X-ray radiation than the wild-type strain. (Meyer *et al.*, 2007) Therefore, it is possible that some

modification had occurred in the Δ*pyrG*Δ*kusA* strain which had weakened the performance of homologous recombination. With the Δ*pyrG* strain transformants were obtained when using homologous recombination, however, it was difficult to obtain a pure strain even after several purification rounds. In some cases, the deletion cassette was apparently integrated to the correct place, but the ORF was still detected. An explanation could be that there is more than one copy of the gene or the stain has become a heterokaryon. If the genes of interest would be essential for the survival for the cell, this would explain the difficulty to delete it (Meyer *et al.*, 2007). However, these genes are not believed to be essential.

When deleting genes with theCRISPR/Cas9 method using a donor DNA, the donor DNA needs sufficiently long flanks. Deleting potential genes related to the galactaric acid pathway using the CRISPR/Cas9 method and a deletion cassette with a selection marker as a donor DNA increased the probability of obtaining the correct strain significantly. In the case of constructing the Δ*39114*Δ*gaaA/C*::*udh* strain using the CRISPR/Cas9, the gene *gaaA/C* was not successfully deleted from the Δ*39114* strain when using the *gpdA-udh* integration cassette with 60 bp flanking regions as a donor DNA. On the other hand, when generating the Δ*gaaA udh* Δ*39114* strain, the *39114* gene was deleted by CRISPR/Cas9 using the deletion cassette with 1.5 kb flaking regions as a donor DNA. It is probable that the 60 bp flanking regions were too short to trigger homologous recombination, because short flanking regions decrease the HR frequency. Flanking regions shorter than 200 bp have shown to reduce HR frequency significantly. Apparently, HR machinery is not able to mediate integration efficiently when the length of the flanking region is too short. Using flanking regions of 1-1.5 kb would give higher HR frequency. (Meyer *et al.*, 2007) The 60 bp flanking regions were decided to use in the deletion, because it was thought that this low frequency problem might be overcome when using CRISPR/Cas9 technology due to the efficiency of CRISPR/Cas9, however, this was not the case.

The deletion using the CRISPR/Cas9 method without a donor DNA harbouring selection marker gave poor results. Despite of efforts to delete the gene *gaaA/C* from the Δ*39114* strain with the CRISPR/Cas9 without donor DNA, the Δ*39114*Δ*gaaA/C* strain was not obtained. In addition, in the construction of the Δ*gaaA udh* Δ*39114* strain, inactivation of the gene *39114* was also not successful. Pressure of selection brought by donor DNA

with selection marker would naturally ease the deletion. However, if the selection was still needed in gene deletion using the CRISPR/Cas9 method, the benefit of using the CRISPR/Cas9 would be lost although it increases the transformation frequency. The gene deletions in this study were conducted with two gRNAs. Using one gRNA should be sufficient to inactivate gene, however, this was not tested in this study. Furthermore, there is a possibility that *in vitro* gRNA does not work as well as *in vivo* gRNA. Thus, using *in vivo* gRNA could be a solution for deleting genes without selection.

8.2 Galactaric acid pathway

Deleting the genes *39114* or *1090836* resulted in a strain with a new phenotype of not utilizing galactaric acid. Moreover, deleting the *1121140* gene weakened the ability of *A. niger* to utilize galactaric acid. The Δ*39114* strain also did not utilize glucaric acid. A hypothesis was that in the catabolic pathway, galactaric acid could be reduced to galacturonic acid and further to galactonic acid, and the rest would be the same as in the galacturonic catabolism. The enzyme 39114 is predicted to be AMP-dependent synthetase and ligase. It was thought that the enzyme 39114 could be the first enzyme in the pathway and reduce the galactaric acid. Therefore, the enzyme 39114 was expressed in *S. cerevisiae* and its activity was tested using galactaric acid as the substrate. Enzyme activity was also tested with glucaric acid, since the Δ*39114* strain also did not utilized the glucaric acid. However, the enzyme 39114 showed no activity on these substrates. Nevertheless, the enzyme 39114 must be related to the galactaric acid pathway, since deletion of this gene resulted in a phenotype of no galactaric acid catabolism. It is possible that the enzyme activity was not detected, because the enzyme assay conditions were not optimal for measuring the activity of the 39114 enzyme or that the yeast was not able to express the gene or translate it to an active protein. Another explanation is that the 39114 enzyme is not the first enzyme of the catabolic pathway. Enzymes 1090836 and 1121140 are also part of the galactaric acid pathway, based on cultivation results. The enzyme 1090836 is predicted to be aldo/keto reductase. The aldo-keto reductase family includes monomeric NADPH-dependent oxidoreductases. Therefore, another hypothesis is that the 1090836 enzyme could catalyse the oxidation of galactaric acid to 2-keto-galactarate with NAD^+ or $NADP^+$ as cofactors. The enzyme 1121140 is predicted to be FAD-dependent oxidoreductase.

Deletion of the *1121140* gene might have resulted in the accumulation of an intermediate, which could not be catabolized further. Enzymes 39114, 1090836 and 1121140 need more investigations to characterize their possible role in the pathway. The activity of the enzymes 1090836 and 1141120 still need measured. The ability of the strains Δ*1090836* and Δ*1121140* to grown on other hexaric acid, such as, glucaric acid has also not been tested, thus, it is unclear whether these genes also block the glucaric acid utilization.

8.3 Gluconokinase

The gene *1162477* was confirmed not to be related to the galactaric acid catabolism. However, the enzyme 1162477 was annotated to be gluconokinase, thus, the enzyme could be involved in the gluconate catabolism. Therefore, the Δ*1162477* strain was assumed not being able to utilize gluconic acid. However, the growth of Δ*1162477* on 2 % gluconic acid plate looked same as compared to the wild-type strain. The assumption might be wrong, because also some *S. cerevisiae* gluconokinase mutants are still able to grow slowly on gluconolactone (van Dijken *et al.*, 2001). In the genome, there is a shorter gene with similar predicted function, which also might have some effect to the gluconic acid consumption.

8.4 Engineering *A. niger* for galactaric acid produciton

The production of galactaric acid from galacturonic acid by *A. niger* was successfully improved in this study. Previously, it has been reported that the *A. niger* Δ*gaaA udh* strain, in which the *gaaA* gene was deleted and the *udh* gene expressed, produced 0.16 g galactaric acid/g galacturonic acid in 4 days. Its parent strain Δ*gaaA* was unable to grow on galacturonic acid, but expressing *udh* restored the ability to grow on galacturonic acid. This showed that the galactaric acid was catabolized by the Δ*gaaA udh* strain. (Mojzita *et al,*. 2010) Deletion of the *gaaA* gene does not block the galacturonic acid pathway completely, whereas deleting the *gaaC* gene does. Therefore, it was thought that the double deletion of the genes *gaaA* and gaaC would be an efficient and secure way to block the galacturonic acid pathway. However, the *gaaA/C* could not be deleted from the Δ*39114* strain. Therefore, the gene *39114* was deleted from the Δ*gaaA udh* strain. The resulting strain, Δ*gaaA udh* Δ*39114,* was able to produce 4.3 ± 0.2

g/l galactaric acid from galacturonic acid with calculated yield of 1.5 ± 0.3 g galactaric acid/g galacturonic acid in 5 days as submerged shake flask fermentation, which is higher than its parent strain Δ*gaaA udh* which produced at best 1.2 ± 0.02 g/l galactaric acid with calculated yield of 0.067 ± 0.001 g galactaric acid/g galacturonic acid. The calculated yield obtained by the Δ*gaaA udh* Δ*39114* strain is higher than the theoretical yield of 1.08 g galactaric acid/g galacturonic acid, which may be due to evaporation during the cultivation. Effect of the evaporation should have been taken into account in order to confirm the actual yield. Nevertheless, the actual yield is believed to be very close to the theoretical yield or even be the same.

8.5 Comparison with other galactaric acid producing strains

The Δ*gaaA udh* Δ*39114* strain is superior to other galactaric acid producing strains. For example, the Δ*gaaA udh* Δ*39114* strain produced same amount of galactaric acid in shorter time than *T. reesei*. With *T. reesei* QM6a *Δgar 1 udh* the obtained yield was 1.08 g galactaric acid/g galacturonic acid in 211 h (Mojzita *et al*., 2010) whereas the Δ*gaaA udh* Δ*39114* strain could do the same in 120 h. The Δ*gaaA udh* Δ*39114* strain can also produce more galactaric acid than *E.* coli. With *E. coli* BL21(DE3) Δ*uxaC* Δ*garD udh* the yield was 0.73 g galactaric acid/g galacturonic acid in 48 h (Zhang *et al.*, 2016) whereas the Δ*gaaA udh* Δ*39114* strain produced 1.4 ± 0.1 g galactaric acid/g galacturonic in 48 h. *A. niger* produces more hydrolytic enzymes than *T. reesei*, which is an advantage when pectin-rich residues are wanted to use as a substrate. Furthermore, using *A. niger* has a clear advantage compared to using *E. coli,* because *E. coli* does not naturally produce hydrolytic enzymes and the substrate would need a separate enzyme treatment.

8.6 Physiology

Additional carbon source is needed for growth to improve galacturonic acid conversion in the Δ*gaaA udh* Δ*39114* strain. Lowering initial cultivation pH from 5 to 3 decreased the galactaric acid production. The same pH-dependent production was also observed by Mojzita *et al.* (2010) and Zhang *et al.* (2016). An explanation is that optimal pH and temperature for *A. tumefaciens* Udh are pH 8.0 and 35 °C (Zhang *et al.*, 2016).

In this study, it was also shown that the strain Δ*gaaA udh* Δ*39114* could produce 16.7 ± 0.9 mg galactaric acid directly from 2 g (1.87 g DM) CPW in SSF, and when the nitrogen was also supplemented the production increased to 28.9 ± 1.1 mg galactaric acid corresponding to 8.9 ± 0.5 and 15.4 ± 0.6 mg galactaric acid/g DM CPW, respectively. Furthermore, the galactaric acid production was increased when the fermentation was performed with the SmF. In this case, the production was 3.1 ± 0.1 g/l galactaric acid from 40 g/l (37.4 g/l DM) CPW corresponding to 82.3 ± 3.9 mg galactaric acid/g DM CPW. L-galactonic and L-ascorbic acid have been produced from CPW with engineered *A. niger* (Kuivanen *et al.*, 2014; Kuivanen *et al.*, 2015). The Δ*gaaB* and Δ*gaaB-gaaA* strains were able to produce 223 ± 2 and 221 ± 6 mg L-galactonic acid/g CPW in SSF from CPW, respectively. In SmF, these strains produced 157 ± 3 and 159 ± 3 mg L-galactonic acid/g CPW. (Kuivanen *et al.*, 2014) The Δ*gaaB* strain has also been further engineered to produce L-ascorbic acid from CPW. The obtained yield with the Δ*gaaB-Eg-Mg (PgaaA/C)* strain was 4.25 mg L-ascorbic acid/g CPW. As a conclusion, engineered *A. niger* have great potentials in producing value-added products from pectin-rich residues. Since *A. niger* is naturally capable of degrading plants, no enzymatic pretreatment to the plant-based residue would be needed, which is the advantage of using *A. niger*.

In the future, it would be interesting to investigate larger scale SSF or SmF by the Δ*gaaA udh* Δ*39114* strain using citrus peel directly as a substrate. The fermentation conditions of the SSF would need optimization, for example, the humidity could be controlled. In this study, the SSF was done without controlling the humidity. In the SmF, the effect of inoculated amount of mycelia to the production could also be investigated. Moreover, starting the SmF by inoculating spores directly to the culture medium instead of inoculating pre-grown mycelia would be interesting. In addition, the effect to galactaric acid production by blocking galactaric acid by deleting the genes *1090836* and *1121140* compared to the deletion of the gene *39114* could also be investigated. Deleting the genes *1090836* or 1121140 alone might not improve the production, but double deletion with *39114* might have a more effective impact.

9 Conclusions

Three genes *39114*, *1090836* and *1121140* were identified to be involved in the galactaric acid catabolic pathway in *A. niger*. Genes *1117792*, *1141260*, *1162477*, *1146483* and *1170646* do not take part in galactaric acid pathway.

An *A. niger* strain with improved galactaric acid production was successfully constructed. Blocking galacturonic acid and galactaric acid catabolism by deleting genes *gaaA* and *39114*, repectively, from *A. niger* and expressing heterologous gene *udh*, resulted in a *A. niger* strain, Δ*gaaA udh* Δ*39114*, which could produce 4.3 ± 0.2 g/l galactaric acid from 20 g/l galacturonic acid in rich medium with a high yield suggesting stoichiometric conversion. In addition when the strain was fermented with CPW, it produced 15.4 ± 0.6 mg galactaric acid/g DM CPW from 2 g (1.87 g DM) CPW in SFF when nitrogen was supplemented and 82.3 ± 3.9 mg galactaric acid/g DM CPW from 40 g/l (37.4 g/l DM) CPW in SmF when using *A. nidulans* defined minimal medium.

CRISPR/Cas9 technology improves gene deletion in *A. niger* and quicken the purification of mutant strains. According to the best of my knowledge, this is the first report of using *in vitro* gRNA successfully in *A. niger* for gene deletion using CRISPR/Cas9 genome editing tool.

References

Acree, S. F., Method of converting wood into mucic acid, U.S. patent US 1816137 A, 1931.

Andersen, M.R., Salazar, M.P., Schaap, P.J., Van De Vondervoort, P.J.I., Culley, D., Thykaer, J., Frisvad, J.C., Nielsen, K.F., Albang, R., Albermann, K., Berka, R.M., Braus, G.H., Braus-Stromeyer, S.A., Corrochano, L.M., Dai, Z., Van Dijck, P.W.M., Hofmann, G., Lasure, L.L., Magnuson, J.K., Menke, H., Meijer, M., Meijer, S.L., Nielsen, J.B., Nielsen, M.L., Van Ooyen, A.J.J., Pel, H.J., Poulsen, L., Samson, R.A., Stam, H., Tsang, A., Van Den Brink, J.M., Atkins, A., Aerts, A., Shapiro, H., Pangilinan, J., Salamov, A., Lou, Y., Lindquist, E., Lucas, S., Grimwood, J., Grigoriev, I.V., Kubicek, C.P., Martinez, D., Van Peij, N.N.M.E., Roubos, J.A., Nielsen, J. & Baker, S.E. 2011, "Comparative genomics of citric-acid-producing Aspergillus niger ATCC 1015 versus enzyme-producing CBS 513.88", *Genome research,* vol. 21, no. 6, pp. 885-897.

Andberg, M., Maaheimo, H., Boer, H., Penttilä, M., Koivula, A. & Richard, P. 2012, "Characterization of a novel Agrobacterium tumefaciens galactarolactone cycloisomerase enzyme for direct conversion of D-galactarolactone to 3-deoxy-2-keto-L-threo-hexarate", *Journal of Biological Chemistry,* vol. 287, no. 21, pp. 17662-17671.

Anonymous, Green FDCA Production, Intratec Solutions LLC, edited Chemical Engineering, 2015.

Anonymous, Commercial Production of Mucic Acid, Chemical and Metallurgical Engineering **26**:1118 (1922).

Arazoe, T., Ogawa, T., Miyoshi, K., Yamato, T., Ohsato, S., Sakuma, T., Yamamoto, T., Arie, T. & Kuwata, S. 2015a, "Tailor-made TALEN system for highly efficient targeted gene replacement in the rice blast fungus", *Biotechnology and bioengineering,* vol. 112, no. 7, pp. 1335-1342.

Arazoe, T., Miyoshi, K., Yamato, T., Ogawa, T., Ohsato, S., Arie, T. & Kuwata, S. 2015b, "Tailor-made CRISPR/Cas system for highly efficient targeted gene replacement in the rice blast fungus", *Biotechnology and bioengineering,* vol. 112, no. 12, pp. 2543-

2549.Dave K., Ahuja M., Jayashri T.N., Sirola R.B. & Punekar N.S. 2012, "A novel selectable marker based on Aspergillus niger arginase expression", *Enzyme and microbial technology,* vol. 51, no. 1, pp. 53-58.

Assadi, M.M. & Jahangiri, M.R. 2001, "Textile wastewater treatment by Aspergillus niger", *Desalination,* vol. 141, no. 1, pp. 1-6.

Barrangou, R. & Marraffini, L.A. 2014, "CRISPR-cas systems: Prokaryotes upgrade to adaptive immunity", *Molecular cell,* vol. 54, no. 2, pp. 234-244.

Barratt, R.W., Johnson, G.B. & Ogata, W.N. 1965, "Wild-type and mutant stocks of Aspergillus nidulans.", *Genetics,* vol. 52, no. 1, pp. 233-246.

Boer, H., Maaheimo, H., Koivula, A., Penttilä, M. & Richard, P. 2010, "Identification in Agrobacterium tumefaciens of the D-galacturonic acid dehydrogenase gene", *Applied Microbiology and Biotechnology,* vol. 86, no. 3, pp. 901-909.

Bouvier, J.T., Groninger-Poe, F.P., Vetting, M., Almo, S.C. & Gerlt, J.A. 2014, "Galactaro δ-lactone isomerase: Lactone isomerization by a member of the amidohydrolase superfamily", *Biochemistry,* vol. 53, no. 4, pp. 614-616.

Carvalho, N.D., Arentshorst, M., Jin Kwon, M., Meyer, V. & Ram, A.F. 2010, "Expanding the ku70 toolbox for filamentous fungi: establishment of complementation vectors and recipient strains for advanced gene analyses.", *Applied Microbiology and Biotechnology,* vol. 87, no. 4, pp. 1463-1473.

Colot, H.V., Park, G., Turner, G.E., Ringelberg, C., Crew, C.M., Litvinkova, L., Weiss, R.L., Borkovich, K.A. & Dunlap, J.C. 2006, "A high-throughput gene knockout procedure for Neurospora reveals functions for multiple transcription factors", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 103, no. 27, pp. 10352-10357.

Delmas, S., Llanos, A., Parrou, J.-., Kokolski, M., Pullan, S.T., Shunburne, L. & Archer, D.B. 2014, "Development of an unmarked gene deletion system for the filamentous fungi Aspergillus niger and Talaromyces versatilis", *Applied and Environmental Microbiology,* vol. 80, no. 11, pp. 3484-3487.

Dhillon, G.S., Brar, S.K., Verma, M. & Tyagi, R.D. 2010, "Recent Advances in Citric Acid Bio-production and Recovery", *Food and Bioprocess Technology,* vol. 4, no. 4, pp. 505-529.

Doran, J.B., Cripe, J., Sutton, M. & Foster, B. 2000, "Fermentations of pectin-rich biomass with recombinant bacteria to produce fuel ethanol", *Applied Biochemistry and Biotechnology - Part A Enzyme Engineering and Biotechnology,* vol. 84-86, pp. 141-152.

Doran-Peterson, J., Cook, D.M. & Brandon, S.K. 2008, "Microbial conversion of sugars from plant biomass to lactic acid or ethanol", *Plant Journal,* vol. 54, no. 4, pp. 582-592.

Dudášová Z., Dudáš A. & Chovanec M. 2004, "Non-homologous end-joining factors of Saccharomyces cerevisiae", *FEMS microbiology reviews,* vol. 28, no. 5, pp. 581-601.

Edwards, M.C. & Doran-Peterson, J. 2012, "Pectin-rich biomass as feedstock for fuel ethanol production", *Applied Microbiology and Biotechnology,* vol. 95, no. 3, pp. 565-575.

Fleißner, A. & Dersch, P. 2010, "Expression and export: Recombinant protein production systems for Aspergillus", *Applied Microbiology and Biotechnology,* vol. 87, no. 4, pp. 1255-1270.

Fraefel, W.; Lichti, H. F.; Brunetti, M. , Process for the preparation of new organ transplants, US Patent, 4383832, 1983; *Chem. Abstr.* **1983**, *99*,p43590d.

Fuller K.K., Chen S., Loros J.J. & Dunlap J.C. 2015, "Development of the CRISPR/Cas9 system for targeted gene disruption in Aspergillus fumigatus", *Eukaryotic Cell,* vol. 14, no. 11, pp. 1073-1080.

Groninger-Poe, F.P., Bouvier, J.T., Vetting, M.W., Kalyanaraman, C., Kumar, R., Almo, S.C., Jacobson, M.P. & Gerlt, J.A. 2014, "Evolution of enzymatic activities in the enolase superfamily: Galactarate dehydratase III from agrobacterium tumefaciens C58", *Biochemistry,* vol. 53, no. 25, pp. 4192-4203.

Hsu, P.D., Lander, E.S. & Zhang, F. 2014, "Development and applications of CRISPR-Cas9 for genome engineering", *Cell,* vol. 157, no. 6, pp. 1262-1278.

Joung, J.K. & Sander, J.D. 2013, "TALENs: A widely applicable technology for targeted genome editing", *Nature Reviews Molecular Cell Biology,* vol. 14, no. 1, pp. 49-55.

Katayama T., Tanaka Y., Okabe T., Nakamura H., Fujii W., Kitamoto K. & Maruyama J.-I. 2015, "Development of a genome editing technique using the CRISPR/Cas9 system in the industrial filamentous fungus Aspergillus oryzae", *Biotechnology Letters,* pp. 1-6.

Kawai S., Hashimoto W. & Murata K. 2010, "Transformation of Saccharomyces cerevisiae and other fungi: methods and possible underlying mechanism.", *Bioengineered bugs,* vol. 1, no. 6, pp. 395-403.

Kiely, D.E., Chen, L. & Lin, T.-. 2000, "Synthetic polyhydroxypolyamides from galactaric, xylaric, D-glucaric, and D-mannaric acids and alkylenediamine monomers - some comparisons", *Journal of Polymer Science, Part A: Polymer Chemistry,* vol. 38, no. 3, pp. 594-603.

Knuf, C. & Nielsen, J. 2012, "Aspergilli: Systems biology and industrial applications", *Biotechnology Journal,* vol. 7, no. 9, pp. 1147-1155.

Kuivanen, J., Dantas, H., Mojzita, D., Mallmann, E., Biz, A., Krieger, N., Mitchell, D. & Richard, P. 2014, "Conversion of orange peel to L-galactonic acid in a consolidated process using engineered strains of aspergillus niger", *AMB Express,* vol. 4, no. 1, pp. 1-8.

Kuivanen, J., Penttilä, M. & Richard, P. 2015, "Metabolic engineering of the fungal D-galacturonate pathway for L-ascorbic acid production", *Microbial Cell Factories,* vol. 14, no. 1.

Lakshmi Prabha V. and Punekar N.S. 2004, "Genetic transformation in Aspergilli: Tools of the trade", *Indian Journal of Biochemistry and Biophysics,* vol. 41, no. 5, pp. 205-215.

Lenouvel F., van de Vondervoort P.J.I. & Visser J. 2002, "Disruption of the Aspergillus niger argB gene: A tool for transformation", *Current genetics,* vol. 41, no. 6, pp. 425-431.

Lewis, J.C. 1972, "Germination of bacterial spores by calcium chelates of dipicolinic acid analogues.", *Journal of Biological Chemistry,* vol. 247, no. 6, pp. 1861-1868.

Lewkowski, J. 2001, "Synthesis, chemistry and applications of 5-hydroxymethyl-furfural and its derivatives", *Arkivoc,* vol. 2001, no. 1, pp. 17-54.

Li, X., Wu, D., Lu, T., Yi, G., Su, H. & Zhang, Y. 2014, "Highly efficient chemical process to convert mucic acid into adipic acid and DFT studies of the mechanism of the rhenium-catalyzed deoxydehydration", *Angewandte Chemie - International Edition,* vol. 53, no. 16, pp. 4200-4204.

Liu, B., Ren, Y. & Zhang, Z. 2015, "Aerobic oxidation of 5-hydroxymethylfurfural into 2,5-furandicarboxylic acid in water under mild conditions", *Green Chemistry,* vol. 17, no. 3, pp. 1610-1617.

Liu, R., Chen, L., Jiang, Y., Zhou, Z. & Zou, G. 2015, "Efficient genome editing in filamentous fungus *Trichoderma reesei* using the CRISPR/Cas9 system", *Cell Discovery*, vol. 1, pp. 1-11.

Mable, B.K. & Otto, S.P. 1998, "The evolution of life cycles with haploid and diploid phases", *BioEssays,* vol. 20, no. 6, pp. 453-462.

Mancera, M., Roffé, I., Rivas, M., & Galbis, J. A. (2003). New derivatives of D-mannaric and galactaric acids. synthesis of a new stereoregular nylon 66 analog from carbohydrate-based monomers having the D-manno configuration. *Carbohydrate Research, 338*(10), 1115-1119. doi:10.1016/S0008-6215(03)00093-4.

Max, B., Salgado, J.M., Rodríguez, N., Cortés, S., Converti, A. & Domínguez, J.M. 2010, "Biotechnological production of citric acid", *Brazilian Journal of Microbiology,* vol. 41, no. 4, pp. 862-875.

Mehtiö, T., Nurmi, L., Rämö, V., Mikkonen, H. & Harlin, A. 2015, "Synthesis and characterization of copolyanhydrides of carbohydrate-based galactaric acid and adipic acid", *Carbohydrate research,* vol. 402, pp. 102-110.

Meyer V., Arentshorst M., El-Ghezal A., Drews A.-C., Kooistra R., van den Hondel C.A.M.J.J. & Ram A.F.J. 2007, "Highly efficient gene targeting in the Aspergillus niger kusA mutant", *Journal of Biotechnology,* vol. 128, no. 4, pp. 770-775.

Meyer, V., Wu, B. & Ram, A.F.J. 2011, "Aspergillus as a multi-purpose cell factory: Current status and perspectives", *Biotechnology Letters,* vol. 33, no. 3, pp. 469-476.

Mohnen, D. 2008, "Pectin structure and biosynthesis", *Current opinion in plant biology,* vol. 11, no. 3, pp. 266-277.

Mojzita, D., Wiebe, M., Hilditch, S., Boer, H., Penttilä, M. & Richard, P. 2010, "Metabolie engineering of fungal strains for conversion of D-galacturonate to meso-Galactarate", *Applied and Environmental Microbiology,* vol. 76, no. 1, pp. 169-175.

Nabyl, M., James, M. B., and Christian, B., Method for the oxidation of aldehydes, hemiacetals and primary alcohols, U.S. patent US 6498269 B1, 2002.

Nga, B.H., Teo, S.P. & Lim, G. 1975, "The occurrence in nature of a diploid strain of Aspergillus niger", *Journal of general microbiology,* vol. 88, no. 2, pp. 364-366.

Nødvig C.S., Nielsen J.B., Kogle M.E. & Mortensen U.H. 2015, "A CRISPR-Cas9 system for genetic engineering of filamentous fungi", *PLoS ONE,* vol. 10, no. 7.

Ochiai, H. & Yamamoto, T. 2015, "Genome editing using zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs)" in *Targeted Genome Editing Using Site-Specific Nucleases: ZFNs, TALENs, and the CRISPR/Cas9 System*, pp. 3-24.

Ogata, N., Sanui, K. and Kayama, Y. (1977), Copolycondensation of hydroxyl diesters and active diesters with hexamethylenediamine. J. Polym. Sci. Polym. Chem. Ed., 15: 1523–1526. doi: 10.1002/pol.1977.170150627.

Omar, S.A. 2016, "Decolorization of different textile dyes by isolated aspergillus niger", *Journal of Environmental Science and Technology,* vol. 9, no. 1, pp. 149-156.

O'Neill, M.A., Ishii, T., Albersheim, P. & Darvill, A.G. 2004, *Rhamnogalacturonan II: Structure and function of a borate cross-linked cell wall pectic polysaccharide*.

Parkkinen, T., Boer, H., Jänis, J., Andberg, M., Penttilä, M., Koivula, A. & Rouvinen, J. 2011, "Crystal structure of uronate dehydrogenase from Agrobacterium tumefaciens", *Journal of Biological Chemistry,* vol. 286, no. 31, pp. 27294-27300.

Pourbafrani, M., Forgács, G., Horváth, I.S., Niklasson, C. & Taherzadeh, M.J. 2010, "Production of biofuels, limonene and pectin from citrus wastes", *Bioresource technology,* vol. 101, no. 11, pp. 4246-4250.

Price, M.S., Classen, J.J. & Payne, G.A. 2001, "Aspergillus niger absorbs copper and zinc from swine wastewater", *Bioresource technology,* vol. 77, no. 1, pp. 41-49.

Punt, P.J., Van Biezen, N., Conesa, A., Albers, A., Mangnus, J. & Van Den Hondel, C. 2002, "Filamentous fungi as cell factories for heterologous protein production", *Trends in biotechnology,* vol. 20, no. 5, pp. 200-206.

Rakus, J.F., Kalyanaraman, C., Fedorov, A.A., Fedorov, E.V., Mills-Groninger, F.P., Toro, R., Bonanno, J., Bain, K., Sauder, J.M., Burley, S.K., Almo, S.C., Jacobson, M.P. & Gerlt, J.A. 2009, "Computation-facilitated assignment of the function in the enolase superfamily: A regiochemically distinct galactarate dehydratase from Oceanobacillus iheyensis", *Biochemistry,* vol. 48, no. 48, pp. 11546-11558.

Ramachandran, S., Fontanille, P., Pandey, A. & Larroche, C. 2006, "Gluconic acid: Properties, applications and microbial production", *Food Technology and Biotechnology,* vol. 44, no. 2, pp. 185-195.

Richard, P., & Hilditch, S. (2009). D-galacturonic acid catabolism in microorganisms and its biotechnological relevance. *Applied Microbiology and Biotechnology, 82*(4), 597-604. doi:10.1007/s00253-009-1870-6.

Rivas, B., Torrado, A., Torre, P., Converti, A. & Domínguez, J.M. 2008, "Submerged citric acid fermentation on orange peel autohydrolysate", *Journal of Agricultural and Food Chemistry,* vol. 56, no. 7, pp. 2380-2387.

Roukas, T. 2000, "Citric and gluconic acid production from fig by Aspergillus niger using solid-state fermentation", *Journal of Industrial Microbiology and Biotechnology,* vol. 25, no. 6, pp. 298-304.

Sander J.D. & Joung J.K. 2014, "CRISPR-Cas systems for editing, regulating and targeting genomes", *Nature biotechnology,* vol. 32, no. 4, pp. 347-350.

Sauermann, G., Schreiner, V., Doring, T., Siefken, W., Gatermann, C., Carstensen, S., and Biergiesser, H., Cosmetic and dermatological preparations containing carnitine for treating and actively preventing dry skin and other negative alterations in the physiological homeostasis of healthy skin, U.S. patent 20040253282 A1, 2004.

Schuster, E., Dunn-Coleman, N., Frisvad, J. & Van Dijck, P. 2002, "On the safety of Aspergillus niger - A review", *Applied Microbiology and Biotechnology,* vol. 59, no. 4-5, pp. 426-435.

Shiramizu, M. & Toste, F.D. 2013, "Expanding the scope of biomass-derived chemicals through tandem reactions based on oxorhenium-catalyzed deoxydehydration", *Angewandte Chemie - International Edition,* vol. 52, no. 49, pp. 12905-12909.

Singh, O.V., Jain, R.K. & Singh, R.P. 2003, "Gluconic acid production under varying fermentation conditions by Aspergillus niger", *Journal of Chemical Technology and Biotechnology,* vol. 78, no. 2-3, pp. 208-212.

Taguchi, Y., Oishi, A. & Iida, H. 2008, "One-step synthesis of dibutyl furandicarboxylates from galactaric acid", *Chemistry Letters,* vol. 37, no. 1, pp. 50-51.

Torrado, A.M., Cortés, S., Salgado, J.M., Max, B., Rodríguez, N., Bibbins, B.P., Converti, A. & Domínguez, J.M. 2011, "Citric acid production from orange peel wastes by solid-state fermentation", *Brazilian Journal of Microbiology,* vol. 42, no. 1, pp. 394-409.

Van Dijken, J.P., Van Tuijl, A., Luttik, M.A.H., Middelhoven, W.J. & Pronk, J.T. 2002, "Novel pathway for alcoholic fermentation of δ-gluconolactone in the yeast Saccharomyces bulderi", *Journal of Bacteriology,* vol. 184, no. 3, pp. 672-678.

Van De Vyver, S. & Román-Leshkov, Y. 2013, "Emerging catalytic processes for the production of adipic acid", *Catalysis Science and Technology,* vol. 3, no. 6, pp. 1465-1479.

Van Es, D. S., van Haveren, J., Raaijmakers, H. W. C., Franciscis, G. P., van Engelen, M., and Frissen A. E., Catalytic oxidation of uronic acids to aldaric acids, UA patent US 20150065749 A1, 2015.

Van Maris, A.J.A., Abbott, D.A., Bellissimi, E., van den Brink, J., Kuyper, M., Luttik, M.A.H., Wisselink, H.W., Scheffers, W.A., van Dijken, J.P. & Pronk, J.T. 2006, "Alcoholic fermentation of carbon sources in biomass hydrolysates by Saccharomyces cerevisiae: Current status", *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology,* vol. 90, no. 4, pp. 391-418.

Verho, R., Putkonen, M., Londesborough, J., Penttilä, M. & Richard, P. 2004, "A Novel NADH-linked L-Xylulose Reductase in the L-Arabinose Catabolic Pathway of Yeast", *Journal of Biological Chemistry,* vol. 279, no. 15, pp. 14746-14751.

Ward, O.P., Qin, W.M., Dhanjoon, J., Ye, J. & Singh, A. 2005, "Physiology and biotechnology of Aspergillus", *Advances in Applied Microbiology*, vol. 58, pp. 1-75.

Wen, S.Y., Fedorov, A.A., Fedorov, E.V., Almo, S.C. & Gerlt, J.A. 2007, "Evolution of enzymatic activities in the enolase superfamily: L-talarate/galactarate dehydratase from Salmonella typhimurium LT2", *Biochemistry,* vol. 46, no. 33, pp. 9564-9577.

Wu, B., Xu, Y., Bu, Z., Wu, L., Li, B. & Dubois, P. 2014 "Biobased poly(butylene 2,5-furandicarboxylate) and poly(butylene adipate-co-butylene 2,5-furandicarboxylate)s: From synthesis using highly purified 2,5-furandicarboxylic acid to thermo-mechanical properties", *Polymer,* vol. 55, no. 16, pp. 3648-3655.

Yelton, M.M., Hamer, J.E. & Timberlake, W.E. 1984, "Transformation of Aspergillus nidulans by using a trpC plasmid.", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 81, no. 5, pp. 1470-1474.

Zhang, H., Li, X., Su, X., Ang, E. L., Zhang, Y. & Zhao, H. 2016, "Production of adipic acid from sugar beet residue by combined biological and chemical catalysis", *ChemCatChem,* vol. 8.

Appendix 1: Primer lists

Table 1 List of primers used to amplify 5' flanking reagion and 3' flanking region from *Aspergillu niger* ATCC1015 genome for deletion cassette constriction for the genes *39114, 1090836, 1117792, 1141260, 1121140, 1162477, 1146483* and *1170646* (JGI protein ID numbers) .

Primer name	Sequence	Usage
oPEEL-002	CCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGGCGGCCGCGTGAAGACCTCGATTTGACC	For the deletion cassette of ID1162477 (gluconokinase), 5´flank for, flank for EcoRI/BamHI digested B2974
oPEEL-003	CTGGTATAGCCAAACATCGCCAATCACCTCAATCACCCGGCAGAACTTATGTATGACGTC	For the deletion cassette of ID1162477 (gluconokinase), 5´flank rev, flank for XmaI digested pyrG (B4343)
oPEEL-004	GCCATGCGGGCTTGGGACGCCATGTCCGTCGCGTGATAACTCCTGACACCTCTATTGAGC	For the deletion cassette of ID1162477 (gluconokinase), 3´flank for, flank for XmaI digested pyrG (B4343)
oPEEL-005	AGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCGCGGCCGCGCAGATAGACATGCCTGAAC	For the deletion cassette of ID1162477 (gluconokinase), 3´flank rev, flank for EcoRI/BamHI digested B2974
oPEEL-007	CCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGGCGGCCGCCACGCGTTTCCACATCTTCT	For the deletion cassette of ID1146483 (dgdB), 5´flank for, flank for EcoRI/BamHI digested B2974
oPEEL-008	CTGGTATAGCCAAACATCGCCAATCACCTCAATCACCCGGATTAGCAATGTGCTGCTTGC	For the deletion cassette of ID1146483 (dgdB), 5´flank rev, flank for XmaI digested pyrG (B4343)
oPEEL-009	GCCATGCGGGCTTGGGACGCCATGTCCGTCGCGTGATAACGTGATTGCGGAGGTGATCTG	For the deletion cassette of ID1146483 (dgdB), 3´flank for, flank for XmaI digested pyrG (B4343)
oPEEL-010	AGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCGCGGCCGCGAAACTGATCGATCAGTCAC	For the deletion cassette of ID1146483 (dgdB), 3´flank rev, flank for EcoRI/BamHI digested B2974
oPEEL-017	CCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGGCGGCCGCGACAGGGTTGAGCCAGTCTA	5' flank aplification for An ID 1090836 del cassette, For, flank for EcoRI/BamHI dig pPEEL-001
oPEEL-018	CTGGTATAGCCAAACATCGCCAATCACCTCAATCACCCGGTGCGACTAGTTGGGTGTCAC	5' flank aplification for An ID 1090836 del cassette, Rev, flank for pyrG (pPEEL-002)
oPEEL-019	GCCATGCGGGCTTGGGACGCCATGTCCGTCGCGTGATAACCAGGATTGTACAACTTAGTT	3' flank aplification for An ID 1090836 del cassette, For, flank for pyrG (pPEEL-002)
oPEEL-020	AGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCGCGGCCGCGTACATCTCAGGGTAATATC	3' flank aplification for An ID 1090836 del cassette, Rev, flank for EcoRI/BamHI dig pPEEL-001

Appendix 1: Primer lists

oPEEL-024	CCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGGCGGCCGCCTACGAGCCGTAGATTGCGT	5' flank aplification for An ID 1117792 del cassette, For, flank for EcoRI/BamHI dig pPEEL-001
oPEEL-025	CTGGTATAGCCAAACATCGCCAATCACCTCAATCACCCGGATGCATGGATGCATGGATAC	5' flank aplification for An ID 1117792 del cassette, Rev, flank for pyrG (pPEEL-002)
oPEEL-026	GCCATGCGGGCTTGGGACGCCATGTCCGTCGCGTGATAACCAGTCGTGTTAAAGCGATCC	3' flank aplification for An ID 1117792 del cassette, For, flank for pyrG (pPEEL-002)
oPEEL-027	AGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCGCGGCCGCAAGGAGCGCGAGGGAATCAGC	3' flank aplification for An ID 1117792 del cassette, Rev, flank for EcoRI/BamHI dig pPEEL-001
oPEEL-031	CCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGGCGGCCGCGAACCAGGCGCAGCGGATC	5' flank aplification for An ID 1141260 del cassette, For, flank for EcoRI/BamHI dig pPEEL-001
oPEEL-032	CTGGTATAGCCAAACATCGCCAATCACCTCAATCACCCGGCTTTCTTACGACGTTGATCC	5' flank aplification for An ID 1141260 del cassette, Rev, flank for pyrG (pPEEL-002)
oPEEL-033	GCCATGCGGGCTTGGGACGCCATGTCCGTCGCGTGATAACGTCCTGTAGTAGTAGGATAA	3' flank aplification for An ID 1141260 del cassette, For, flank for pyrG (pPEEL-002)
oPEEL-034	AGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCGCGGCCGCGTTGCTATCACAAAGGATTC	3' flank aplification for An ID 1141260 del cassette, Rev, flank for EcoRI/BamHI dig pPEEL-001
oPEEL-036	CCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGGCGGCCGCGTATTGGCTTAACCCACCCT	5' flank aplification for An ID 1121140 del cassette, For, flank for EcoRI/BamHI dig pPEEL-001
oPEEL-037	CTGGTATAGCCAAACATCGCCAATCACCTCAATCACCCGGCGAGTGATTCTTCGATTATA	5' flank aplification for An ID 1121140 del cassette, Rev, flank for pyrG (pPEEL-002)
oPEEL-038	GCCATGCGGGCTTGGGACGCCATGTCCGTCGCGTGATAACCAGTAACAGTAATCGTAGCAG	3' flank aplification for An ID 1121140 del cassette, For, flank for pyrG (pPEEL-002)
oPEEL-039	AGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCGCGGCCGCACGGCGCCAATGAGATATGC	3' flank aplification for An ID 1121140 del cassette, Rev, flank for EcoRI/BamHI dig pPEEL-001
oPEEL-040	CCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGGCGGCCGCCTAAGCTTTAGCTACAAGCA	5' flank aplification for An ID39114 (AMP ligase) cassette, For, flank for EcoRI/BamHI dig pPEEL-001
oPEEL-041	CTGGTATAGCCAAACATCGCCAATCACCTCAATCACCCGGCGTTGTTACATATAGAAGCA	5' flank aplification for An ID39114 (AMP ligase) del cassette, Rev, flank for pyrG (pPEEL-002)
oPEEL-042	GCCATGCGGGCTTGGGACGCCATGTCCGTCGCGTGATAACGTGGTCTATGTTGGATAGAT	3' flank aplification for An ID39114 (AMP ligase) del cassette, For, flank for pyrG (pPEEL-002)
oPEEL-043	AGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCGCGGCCGCTCCCAGTCCTCATCATCATC	3' flank aplification for An ID39114 (AMP ligase) del cassette, Rev, flank for EcoRI/BamHI dig pPEEL-001
oPEEL-044	CCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGGCGGCCGCCGACCGATGGACACTCTTGT	5' flank aplification for An ID1170646 cassette, For, flank for EcoRI/BamHI dig pPEEL-001

oPEEL-045	CTGGTATAGCCAAACATCGCCAATCACCTCAATCACCCGGTGTAGGCAAAGGGTGAGAGT	5' flank aplification for An ID1170646 del cassette, Rev, flank for pyrG (pPEEL-002)
oPEEL-046	GCCATGCGGGCTTGGGACGCCATGTCCGTCGCGTGATAACCATTGGCTTCGCGCTGAAAT	3' flank aplification for An ID1170646 del cassette, For, flank for pyrG (pPEEL-002)
oPEEL-047	AGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCGCGGCCGCGCAAAGCAGACGTTAAGCCC	3' flank aplification for An ID 1170646 del cassette, Rev, flank for EcoRI/BamHI dig pPEEL-001

Table 2 List of primers used for *Aspergillus niger* colony PCR to verify the correct integration of the deletion cassette. Primers are used to amplify the 5' flanking region of the cassette around 1.5 kb (the primers anneal to the beginning of *pyrG* selection marker and outside 5' flank), part of the open reading frame (ORF) of the gene around 700 bp and 3' flanking region around 1.6-1.7 kb (the primers anneal to the end of *pyrG* selection marker and outside 3' flank).

Primer name	Sequence	Usage
oPEEL-001	AGCTGGTATAGCCAAACATC	Rev primer for A. niger pyrG, anneals at promoter region
oPEEL-006	GTAACACAAGAGAGGCAGAC	For primer for checking 1162477 deletion in A. niger, should be used with oPEEL-001
oPEEL-011	GCTAATACGTGGTATGTATG	For primer for checking ID1146483 (dgdB) deletion in A. niger, should be used with oPEEL-001
oPEEL-021	GAAGGCAGGATTGGAGAAGG	For primer upstream from An ID 1090836 del cas 5'flank in An gDNA, can be used for A. niger colony PCR
oPEEL-028	CACGTGCTACGCCAGGTAC	For primer upstream from An ID 1117792 del cas 5'flank in An gDNA, can be used for A. niger colony PCR
oPEEL-029	GTAGCATGGAGGTAAAGTAT	For primer for checking ID39114 (AMP ligase) deletion in A. niger, should be used with oPEEL-001
oPEEL-030	CAGCTATGCTACAGTATATC	For primer for checking ID1121140 (FAD enzyme) deletion in A. niger, should be used with oPEEL-001
oPEEL-035	CAAGCTACTATCGAGCACTC	For primer for checking ID1141260 deletion in A. niger, should be used with oPEEL-001
oPEEL-048	GTAGACTACCAAGTCGTAGT	For primer for checking 1170646 deletion in A. niger, should be used with oPEEL-001 or -023
oPEEL-059	TTGAGCTCGTGTGTCTGGAC	For checking mixed population. Amplify ORF ID 39114, FOR
oPEEL-060	ACAGCGGCTAGATAACGAGC	For checking mixed population. Amplify ORF ID 39114, REV
oPEEL-061	GATACATGGGCAGCGATGGA	For checking mixed population. Amplify ORF ID 1090836, FOR
oPEEL-062	CTCTCCTCACCCACCTCTGA	For checking mixed population. Amplify ORF ID 1090836, REV
oPEEL-063	ATGACCATCACCGAACCCAC	For checking mixed population. Amplify ORF ID 1117792, FOR

Appendix 1: Primer lists

oPEEL-064	GGCTCGCGCAAACTGAATAG	For checking mixed population. Amplify ORF ID 1117792, REV
oPEEL-065	CAAGGTCATCGTCGCGGATA	For checking mixed population. Amplify ORF ID 1141260, FOR
oPEEL-066	CCAACGGCACATTTCCGATG	For checking mixed population. Amplify ORF ID 1141260, REV
oPEEL-067	GCATAGCAGTGCTTCTTCGC	For checking mixed population. Amplify ORF ID 1121140, FOR
oPEEL-068	CTCCGTTGTTCACTCCGTCA	For checking mixed population. Amplify ORF ID 1121140, REV
oPEEL-069	AGTAGATTCCACACGCGGTC	For checking mixed population. Amplify ORF ID 1162477, FOR
oPEEL-070	GCCAACGTATCCCTCACCAA	For checking mixed population. Amplify ORF ID 1162477, REV
oPEEL-071	GCTGGCTTTTCGTCAAGGTG	For checking mixed population. Amplify ORF ID 1146483, FOR
oPEEL-072	GCTTGATAGCTTCGGGGTGT	For checking mixed population. Amplify ORF ID 1146483, REV
oPEEL-073	GCCCATAATCCTCCACCTCG	For checking mixed population. Amplify ORF ID 1170646, FOR
oPEEL-074	CAGTCAACGGACAATGCACC	For checking mixed population. Amplify ORF ID 1170646, REV
oPEEL-088	CCCCTATACCCGTCTGTTTG	For primer for checking deletion in A. niger, binds to the end of pyrG
oPEEL-089	CCATGGCATCCTCGAGCTCC	Rev primer for checking deletion ID 39114 in A. niger, used with oPEEL-088, ~1.7 kb
oPEEL-090	GTTGTTCATCGATTCCCCCG	Rev primer for checking deletion ID 1090836 in A. niger, used with oPEEL-088, ~1.7 kb
oPEEL-091	CATCCTCGACTGCAGCAATG	Rev primer for checking deletion ID 1117792 in A. niger, used with oPEEL-088, ~1.7 kb
oPEEL-092	CGTACTGTCAGAGCAACCGA	Rev primer for checking deletion ID 1141260 in A. niger, used with oPEEL-088, ~1.7 kb
oPEEL-093	CCGGGGGTAATAGTAGTCGC	Rev primer for checking deletion ID 1121140 in A. niger, used with oPEEL-088, ~1.7 kb
oPEEL-094	CATACTGGATTCTGTGGGCC	Rev primer for checking deletion ID 1162477 in A. niger, used with oPEEL-088, ~1.6 kb
oPEEL-095	GTGCTCTTCTAGGTTCCGAG	Rev primer for checking deletion ID 1146483 in A. niger, used with oPEEL-088, ~ 1.7 kb
oPEEL-096	GGCGGCACCGTCCGTGTGTG	Rev primer for checking deletion ID 1170646 in A. niger, used with oPEEL-088, ~1.7 kb

Appendix 1: Primer lists

Table 3 List of primers to amplify open reading frame (ORF) of the gene for expression in heterologous host and to sequence the ORF.

Primer name	Sequence	Usage
oPEEL-049	CTAGAATTCAACAAAATGCTCTCCACAATCGCACC	Amplification of An ID 39114 ORF (AMP ligase) with EcoRI site, for
oPEEL-050	CTTGGATCCTCAAACCCACGCAACGTCCC	Amplification of An ID 39114 ORF (AMP ligase) with BamHI site, rev
oPEEL-075	GCTCGGCTGCTGTAACAGGG	Sequencing primer 1 ORF ID 39114 (AMP ligase)
oPEEL-076	GCGAGTACCTGAGCTAGTCA	Sequencing primer 2 ORF ID 39114 (AMP ligase)
oPEEL-077	GCGATTCGAAGAACGCTTTC	Sequencing primer 3 ORF ID 39114 (AMP ligase)
oPEEL-078	GCGTTCCATACTCGTCCAGA	Sequencing primer 4 ORF ID 39114 (AMP ligase)
oPEEL-079	CCCGCAACCTTCAACTTCTG	Sequencing primer 5 ORF ID 39114 (AMP ligase)
oPEEL-145	AAAAAGAATTCAACAAAATGGCATTAAACCGTACCTT	Amplification of ID 1090836 ORF with EcoRI site, for
oPEEL-146	TTTGGATCCCTAAGCCTTCGCCTTCAGCA	Amplification of ID 1090836 ORF with BamHI site, rev
oPEEL-151	AAAAAGAATTCAACAAAATGGGCTCATCATTATCCTC	Amplification of ID 1121140 ORF with EcoRI site, for
oPEEL-152	TTTGGATCCTCACACCTCTTGCTTCACAT	Amplification of ID 1121140 ORF with BamHI site, rev
oPEEL-164	GCGCTGGGGAATTACACCTC	Sequencing primer 1 ORF ID 1121140
oPEEL-165	GCCCAGACTACCAGCGGGCC	Sequencing primer 2 ORF ID 1121140
OK19	CTATTTTCCCTTCTTACG	Sequencing primer, binding p2159, 102 bp upstream EcoRI restriction site, for
OK20	CGTTCATTGTTCCTTATTC	Sequencing primer, binding p2159, 27 bp downstream BamHI restriction site, rev

Table 4 List of oligonucleotides or primers used for constructing galactaric acid producing *Aspergillus niger* strain. Guide RNA (gRNA) were used in CRISPR/Cas9 system to create double-stranded beak to the genes *gaaA* and *gaaC*. Gene *udh* was integrated to *gaaA* and *gaaC* locus in *Aspergillus niger*. Primers oPEEL-086 and -087 were used for colony PCR to confirm the deletion of *gaaA* and *gaaCI* and integration of *uhd*.

Primer name	Sequence	Usage
oPEEL-051	GCTGCAAGGCGATTAAGTTG	Amplification of gpdAp-udh-hph from B6042 for the gaaA/C integration cassette, FOR
oPEEL-052	CAATACGCAAACCGCCTCTC	Amplification of gpdAp-udh-hph from B6042 for the gaaA/C integration cassette, REV
oPEEL-053	CCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGGCGGCCGCACTTGCACTGGAAGAGC GTT	5'flank amplification for udh gaaA/C integration cassette, flank for EcoRI/BamHI dig pPEEL-001, FOR
oPEEL-054	GGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCATCTTCAACCTCTGGAGCGC	5'flank amplification for udh gaaA/C integration cassette, flank for udh hph cassette, REV
oPEEL-055	TGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGAGAGAAGCTTGTATGTAATTACTTT GT	3'flank amplification for udh gaaA/C integration cassette, flank for udh hph cassette, FOR
oPEEL-056	AGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCGCGGCCGCGTTAATTCCCTTAGCGC GGCT	3'flank amplification for udh gaaA/C integration cassette, flank for EcoRI/BamHI dig pPEEL-001, REV
oPEEL-080	TAATACGACTCACTATAGTGTAATTACTTTGTTATGTA	gaaA gRNA for
oPEEL-081	TTCTAGCTCTAAAACTACATAACAAAGTAATTACA	gaaA gRNA rev
oPEEL-082	TAATACGAGTCACTATAGTCATAATCTATATGATACAA	gaaC gRNA for
oPEEL-083	TTCTAGCTCTAAAACTTGTATCATATAGATTATGA	gaaC gRNA rev
oPEEL-084	CGACCCCACTGATTTCAGTTCTACTGCAGATACACACACGCCTCGGAAACCCTTCTATTCCTGTACA GTGACCGGTGACT	udh with gaaA/C flanks for, fragment is ~2.5kb
oPEEL-085	TTGTCGCAAACCTGCATGTTCACCGCAACAGATGGACCAGAATCATAAACGGTCATACAGGTGGA GATGTGGAGTGGGCG	udh with gaaA/C flanks rev
oPEEL-086	CTATCAACAACCCGCCCAAG	5´out gaaA/C del for
oPEEL-087	AGCACAGCTGCAGCATTAAG	3' out gaaA/C del rev

Table 5 List of oligonucleotides for synthesizing guide RNA (gRNA) needed in CRISPR/Cas9 system *in vitro* using GeneArt™ Precision gRNA Synthesis Kit (ThermoFisher Scientific). CRISPR/Cas9 was used to delete genes. Sequences TAATACGACTCACTATAG (forward) and TTCTAGCTCTAAAAC (reverse) were added to the primers according to the kit.

Primer name	Sequence	Usage
oPEEL-097	TAATACGACTCACTATAGTATCACCTCCGCATAACCAT	gRNA 5' 1090836 for
oPEEL-098	TTCTAGCTCTAAAACATGGTTATGCGGAGGTGATA	gRNA 5' 1090836 rev
oPEEL-099	TAATACGACTCACTATAGAGTTCAGTCTCTGTGGAGGA	gRNA 3' 1090836 for
oPEEL-100	TTCTAGCTCTAAAACTCCTCCACAGAGACTGAACT	gRNA 3' 1090836 rev
oPEEL-101	TAATACGACTCACTATAGATGCATCCATTTTAGAGCCA	gRNA 5' 1117792 for
oPEEL-102	TTCTAGCTCTAAAACTGGCTCTAAAATGGATGCAT	gRNA 5' 1117792 rev
oPEEL-103	TAATACGACTCACTATAGTCTGGAAGTCGCTGTGCCCT	gRNA 3' 1117792 for
oPEEL-104	TTCTAGCTCTAAAACAGGGCACAGCGACTTCCAGA	gRNA 3' 1117792 rev
oPEEL-105	TAATACGACTCACTATAGTATCGCCAGAACAAAAAGCG	gRNA 5' 1141260 for
oPEEL-106	TTCTAGCTCTAAAACCGCTTTTTGTTCTGGCGATA	gRNA 5' 1141260 rev
oPEEL-107	TAATACGACTCACTATAGAGACCGATCATTATTGACGA	gRNA 3' 1141260 for
oPEEL-108	TTCTAGCTCTAAAACTCGTCAATAATGATCGGTCT	gRNA 3' 1141260 rev
oPEEL-109	TAATACGACTCACTATAGTCGTGTCTGACATTCCACAA	gRNA 5' 1121140 for
oPEEL-110	TTCTAGCTCTAAAACTTGTGGAATGTCAGACACGA	gRNA 5' 1121140 rev
oPEEL-111	TAATACGACTCACTATAGACAGAGTTCTATTCACGGGT	gRNA 3' 1121140 for
oPEEL-112	TTCTAGCTCTAAAACACCCGTGAATAGAACTCTGT	gRNA 3' 1121140 rev
oPEEL-113	TAATACGACTCACTATAGAGGACGTCATACATAAGTTC	gRNA 5' 1162477 for
oPEEL-114	TTCTAGCTCTAAAACGAACTTATGTATGACGTCCT	gRNA 5' 1162477 rev
oPEEL-115	TAATACGACTCACTATAGATACACTGGGTTGGCTTGCT	gRNA 3' 1162477 for
oPEEL-116	TTCTAGCTCTAAAACAGCAAGCCAACCCAGTGTAT	gRNA 3' 1162477 rev
oPEEL-117	TAATACGACTCACTATAGTGTTACCCCCACGCGGGGTA	gRNA 5' 1146483 for
oPEEL-118	TTCTAGCTCTAAAACTACCCCGCGTGGGGGTAACA	gRNA 5' 1146483 rev
oPEEL-119	TAATACGACTCACTATAGAGACTTGTAGGCCAGGATGT	gRNA 3' 1146483 for

Appendix 1: Primer lists

oPEEL-120	TTCTAGCTCTAAAACACATCCTGGCCTACAAGTCT	gRNA 3' 1146483 rev
oPEEL-121	TAATACGACTCACTATAGAGGCAAAGGGTGAGAGTAGT	gRNA 5' 1170646 for
oPEEL-122	TTCTAGCTCTAAAACACTACTCTCACCCTTTGCCT	gRNA 5' 1170646 rev
oPEEL-123	TAATACGACTCACTATAGAGCTCCCCCCTGCCTCCTCG	gRNA 3' 1170646 for
oPEEL-124	TTCTAGCTCTAAAACCGAGGAGGCAGGGGGGAGCT	gRNA 3' 1170646 rev
oPEEL-129	TAATACGACTCACTATAGTTCTGTCTGAGCCCGAGCTG	ID 39114 5' gRNA for, cut ~541 bp
oPEEL-130	TTCTAGCTCTAAAACCAGCTCGGGCTCAGACAGAA	ID 39114 5' gRNA rev, cut ~541 bp
oPEEL-131	TAATACGACTCACTATAGTCCACGATGCCCTACACACC	ID 39114 3' gRNA for, cut ~541 bp
oPEEL-132	TTCTAGCTCTAAAACGGTGTGTAGGGCATCGTGGA	ID 39114 3' gRNA rev, cut ~541 bp

Appendix 2: Sequence of genes related to galactaric acid catabolic pathway

JGI ID 39114 IPR000873: AMP dependent synthetase and ligase (3246 bp)

ATGCTCTCCACAATCGCACCACAACCACCCACGCTACAAGAGCCTCTGTCTAAAGACGATCATATTCTTCCTTTA
CAATCTTACGAACCCTCGACCATTGACGAGCTCGTTCGTCAGCGTGCATCACTTGGAGCCGCACAACCCATCAT
TTCTTACCCGCGGACCGGCATCGAATATGTGGACTATCCACTGCAGCAACTGGATGTCTTCGCATTCCGCGTCT
CCAAGGTGTTATCAGACCGCATTCCACCTCGGAAGTCGTCGGCTGAAACTCCCAAAGTCATCGCTTTACTGGGA
CCTTCAGATCTCAACTACCTGGTGATGCTGCTTTCACTGGCCAAATTGAGTCACAGTGGACTGCTCTTATCGACC
AGGATCTCCATAGATGCTTATGTCTCCCTGCTGGAGAGAACCGGGTCTCGACATGTTTTCATTCACAGCTCTTTC
CGAGATACCGCCGAGGAAATCAAGAAGCGAGTACCTGAGCTAGTCATTGATGAGATTCCAACTGAGGAAAAC
TATCACTATCCCATCACGGAATATGTCGACACGAATCTGGTTCCGCATCTCGACCCGAAGATTGAATCCAAGCA
CATTGCATGGATTATTCACTCCAGTGGATCGACTGGGCTTCCGAAACCTATCTTTCATACGCAATCTGCTGCGCT
GAAGAACTACTCCGGACATATGAACATGTCTGGGTTTGTCACTCTGCCTCTTTATCACAACCATGGAATCAGTT
GTCTGTTCCGGACGATTCACGCTAGCAAACAGCTGCACCTGTATAATGCCAATCTCCCATTGACAAGGCAGTAC
TTGCTCGAGATCATGGGGTCCAATTCGTTTGAAGTTTTCTACGGTGTCCCCTATGCGCTGAAACTGTTGGCTGA
GACACGAGAGGGAATATCCGCTCTGGCTAAGCTAAAGGCCGTCATGTTTGGAGGCTCCGCATGCCCGGACTCT
CTGGGTAACTTGCTGGTCGAAAATGATGTTCATCTCATCAGCCACTACGGATCGTAAGTGGATTCTGCAATGTA
TTTCGGGTACTTTACTAACATATATACAGAACTGAGACCGGCCAGCTTATGATGTCGACCAGGCCTCGTGACGA
CAAGGGATGGGACTGGCTCCGTCCATCTGACACCGTCAAAAGGTTTTTGCGATTCGAAGAACGCTTTCCGGGA
GTATTTGAGCTCGTGTGTCTGGACGGCTGGCCTTCCAAGGTCATGACCAACCGACCAGACGGCTCGTACGCCA
CCAAGGACCTGTTTGTCAAGCATCCCACTATGGAAGCATACAAGTACTATGCTCGGCTAGATGATACAATCGTT
CTATATAACGGAGAGAAGGTCAACCCACTTGACCTGGAAGGACGAGTCCGACAGCGCAGCACTGTTGCCGAA
GCCATTGCATTTGGAGCAGGCAAAGCCCACATCGGGCTGGCTGTCATCCGTGCCCCTGGCACCGAGTCACTCT
CCGATGAGGATATCATTGACAGTATTTGGCCAGCTGTTGAGAAGGCTCACGAGGCACTACCTGCGTTCGGACA
GCTCTCAAAGAACATGGTTCGAGTGTTGCCTGCGGACACTCCCTATCCTCGGACTGATAAGGGCACCATCATCC
GGCAAGCATTTTACAAGAACTTCCAGCCGCTGATTGAAGAGGTTTACGCCGCTGTGGACGCCATGACGGGCAC
GCTGGTTCTGTCTGAGCCCGAGCTGAGGGACTTCCTCAAGAAGCAGCTTCTCCAGATTTTGCCCCTCAAGGATT
CCAATCTGCTGACCGACGATGCTGATTTCTTCTCCCTGGGCATGGACTCTCTGCAGGCTAGTCAGCTGCGTTCC
ATACTCGTCCAGAACTTGGACACCAAGGGCCATCAACTGGGACTTAATATTGCCTTCGAGCAGCCCACTATTTC
CTTGCTCGCTCGTTATCTAGCCGCTGTGCAGTCCGGAGAAGCCCTTCCAGGTTCTCAGCCAATCCATGAACAAA
TGCGCGCTCTCATCTCCCAATTCAGCCACTTTGAACCACACGTTCCCCATTCTAATGAGCTTCCTGGTCGTTATG
TTGTGAGTATCCCTCAACATCCAGGTCAGTACACACTGCTGACTATTCTAGGTTCTCACCGGTGCCACTGGCTCC
CTAGGCAGCCATATCGCGCACCAACTCGCACAGAACTCTTCCGTGAACAAAGTCTACTGCCTCATCCGCGCCTC
CTCCCCAATTGAAGCCTACAAGCGCCTCCACGATGCCCTACACACCCGGTACGTCTACAGCCCCCTCTCTTCTTC
CTCCAAAGCCAAGCTCATCGCCCTCCCCGCCCCAACCCTATCCCATCCAACTCTCTCCCTCCCTGAGGAAACCTA
CAACACCCTCCTCACCGAAACCACTGACATCATCCACTGCGCCTGGCCCGTCAACTTCAACCTCCAGCTCAGCA
GCTTGGCACAAGACACCCTCCCCACCCTACACAATCTTCTCTCTCTAGCCCTGAAAGCCCAGCGCCCCGAACCC
GCAACCTTCAACTTCTGCTCCTCCGTCAGCAGCGTCGTCAACAGCGCCGTCTCCCCCATCCCCGAAACTCTCCCT
GAGTCCCTCACCGCTGCACAATCAATGGGTTATGCCCAATCCAAGCTCATCGCAGAACATATCTGCGCCAACGC
CACGCCCTACCTTGATGCGCGGGTCCTCCGTATCGGGCAGATCATCGGCGATACGAAGCACGGGGGTGTGGAA
TGCCACGGAGGCAATTCCGTTGATGCTTCGGGCGGCGGTTACAGTGGGGGCGTTGCCGAGGCTGGATGAGA
GGATGCGGTGGGTGCCGGTTGATGTTGTTGCTGCAGCTGTGATGGATATTACTCTTCACAAGGAGGAGCAAG
GGCTGGAAAGGAAGAAGGGGGCGGATGATGTGGAGGTGTATAACATCTTGAACCCGTATAGCTTCCATTGGA
CAAAGGATTTGCTTCCTGCTCTAAGGGCGGCCGGGTTTGGATTTGAGGATATGGAGTTCGCGGAGTGGATTAA
GCGGGTTAAGGATCTGGCAGATCCGGAGAGGAATCCGCCTGTTAAGTTGGTGGGGTTCTGGGAGGGGAAGT
ATGGAAGTGCGAAGCCATTTAGGGGATTGGAGTTTGTCACGGAGAAGGCGAGGGAGAGGGCTGAGGGGTTG
AGGGAGTTGAGTGCGGAAGGGTTGGAAGGGGGACTGGTGGGCAAGATGGTGGAGTGGTTCAGGGACGTTG
CGTGGGTTTGA

NNNNN = exon
NNNNN = intron

JGI ID 1090836 IPR001395: Aldo/keto reductase (1376 bp)

ATGGCATTAAACCGTACCTTCAAGCTCAACACAGGCTACGACATGCCTGCCGTCGGCCTAGGAACATGGGTAT
ATCAATCGCTTTTACGCAAAACATGCCCGCTTCAGATCTGACACAATCAAACAGCAATCGAAGAAAGACGAAG
TCCGCGACGCCGTCATCGCAGCCCTGAAATGCGGCTACCGGCACATTGACGCGGCAGCAGTGTATGGCAACG
AACAGGAAGTGGGAGACGGGATGAGATTGTCTGGGGTTCCGCGGGAGGAGATTTTCGTGAGCCAGCCCTTAA
AACCTAAAGAGAGGGTTATGATATAGGGCTAACTCATGCTGAATGTATAGCTCACAAGCAAGCTATGGAATAC
CCACCACCATCCCGAAAATGTCGAAGAAGCTGTGGATAAATCGCTGGCTGATTTGCAGACTGATTATCTGGAT
CTTTATCTTGTATGTTGACGATGTGTCTACTAGCCGTTGAAGTCCTGCTAAATATTGTACAGATCCACTGGCCCG
TCGCGTTCCGATACTCAACTACCACCATCCAGCCTGTCAATGAACAAACGGGCCTGATTGACGTGGTAGACGTA
CCAATCAAAGATACATGGGCAGCGATGGAGAAACTAGTAGAGAAAGGCAAAGTCCGCTCCATCGGTGTGAGC
AACTTTACTCGCGAGAAGATCGAGGAGTTGTTGAAGACGTGAGTATGCTTGTTATTGCGTATGTATCATAGCT
GACGCTATCTACAGAGCGAAGATTACTCCCGCAGTCAACCAGATCGAGGCTCATCCGTTCCTTCAGCAGAGGG
ATCTTTTGGAGTGGTCGACGCAGAAAGTAAGATACTTTACATCTTCAACAACACGACTATCATACTAATCGAAC
TTTGTCTAGGGCATCGTGGTAGCAGGCTACTCCCCCCTAGGCAACAACATCTACAACATCCCTCGGTATGCAGA
CCGTTTCCAGCTCTTCATATCTCGCAAGAAACTGACAAACTCCCAGCGCCGTCGACGACCCCCTCGTCATCGAA
ACAGCCAAGAAATTAAACAAAACCCCTGCGCAAGTCCTCATCAGCTGGGCCGTCCAGCGCGGAACTGTCGTGC
TCCCCAAGTCCGTCACTCCGGAGAGAATCGAGAGTAATTTCCAAGGTAGTCTTTCTCCCACCCACAACATATGT
CCTCATTGTGCTAATATGGACTATTCAGATTTCGTCCTTCCCGACGATGCATTCTCAACCATTCAATCACTGGAA
CGCCACCAGCGGATGAATTTCCCGGCTAGAATCGGTGTCGATATTTTCTCAGAGGTGGGTGAGGAGAGTGTG
AGGAAGAGTGCGCTTGCTTGGGCGGAGCAGCAGAGGGTGCTGAAGGCGAAGGCTTAG

NNNNN = exon
NNNNN = intron

JGI ID 1121140 IPR006076: FAD dependent oxidoreductase (2353 bp)

ATGGGCTCATCATTATCCTCGGTTTCCAATTCTTCTACCGAAGACATTGTCATAGTAGGCGCTGGCGCCAGTGG
CATAGCAGTGCTTCTTCGCTTGATCGAACATGCTAAGGATGGCAAGAAGATTCCTCCCATCATCGTCGTTGAGA
AAGCTTCACCTCCTGGTCCAGGACTGGCATATTCTGCCGCTTGTACCGGGACCATCCTGAACATGCACACGGAC
ACAATGGGTCTATATTATAACGACCCCAAGCATTTTACGAGATGGAGAAGCGAGCTAGCCAGTGGTCCCTTCC
CGTCCCGAAGCCAATATGGAGAGTACCTGGAGGCGATGTGGTCCGAGATACTCTCGCAAGCCCAACAGATGG
GCTTGGACATCTCAATCATCCAAGACGAAGTATCCGACATTGACCGTCATGACGACAGTACCTTCACCTTGACA
CTGACAGGGGGGCGTCGCCTTGCTGCCCAATCTGTCGTGTTGGCGCTGGGGAATTACACCTCCACTCTCAACA
CGCACCTGCTGAATCAACCGGGCTTCTTTCCCAGCCCGTGGCCAACCTCTCAACTAAAGACCATTCCCGCCGAT
GCCTCGGTCCTCATCATCGGATCAAGGCTCTCCGCTGTTGACGCAGCGCTATTTCTCTCCAAGAATGGCCACAA
GGGCCCGTTGACCTTCATGTCGCGCAGTGGCCGCCTTCCCAAAGTCCAGGGCGAGCCTGAGCCCTACCCGCGC
CGTTACACCCTTCACACCTTGGCCCGATACATCGAATCCAACCCAGCGGACGGTCTTGTCAAGCTCACAACCAC
CCTTATGGATGAAATTGACGGAGTGAACAACGGAGATTGGACCTGGATACAGAAGCATGCCTCGCCGCTCGCT
GAACTACGGGCAGACCTATTCGCGGCACAGGGAGGAAATGTCCATTGGCAAACCGTCCTCCGCCACACGGCC
CCAGTAATCGAGCGTTACTGGCACTGTTTGCCTCTAGAAAGCCAGAAGCTATTCATGGCCAAATTCTTCACTCC
TTGGATGCGATATCGTCACGGAATGCCGGTGCAGAATGCGCAAAAGATACTCAACCTGATGGAAACCTCCCAA
CTTAGTGTCGTCGCAGGAGAGGCTGTCCACTGGGATGAGGGTGAGGGCGCCTTCATCGCCCAGACTACCAGC
GGGCCCATAGAAGCATCCTATGTCATCGAAGCAACTGGACAAGAGAGTCATCTCGATCGCATCCCATCTCCGC
TCGTCCAGTCCGCTGTCCGAAAAGGATTATTCACCCCACACCCGATGGGTGGGGTTGACGTCGACTTCCACACC
CTCTGTACATCAACACCCGGCCTATACACTATGGGCTCGTTGACCCGCGGCACGCACTTCTACGTCAGCGCCAT
CGATCGAACGGCTGCCCATGCGGCTCGCATCGCCGACGCACTGGTCGGGGAGCCCCCTGCCAGACCTGTACAT
ATCGCGATCTTTCTCGGTGCAGACGTGGCCTCGCACCTGATAGCCAGTGACCTGGTCCCCCTGCTTCTGGCCGA
GGGCCATATGCCTTTCCTCTTTTTGTCATCCAGCAAGCAGACACCTTCCTTAGAGGGGTCTGATAGTCGACCTTT
CGACCTGCGCAAACTGGAATTTTTCGAACGCGAGCTTTTCCGCAAGCATCTCTGTCCAAGATTGAAGGAACACA
GTTTCAAAGGCGCACGGCACATGACCGTTGAGCAGATGCAGACGGCATACGGGGTCCTTGTGCAAGAAATAC
CAGACACGAAGGGTGCCTCGGTCGGCAGAATGTTGCAAAAATACTTCATCGACGTCGGCATATCACTGACATG
CGCCGAGGCACCCGACCAGGATGTTATTGCTTACTTATCCTCTTCGTCGCGCCACCTTCTAGCAGTGGATGCCG
GCGTCCTCTCGGCTCCTTGGGACAGCAAGAAAGCTGGAGCCAAGTTTGGGTATACCCTTCGTGAATTTCATGA
AGACGGTAAGTTAGGAGACGTGATTGATCGCCGGACTATTCCCGTCGGGGAGTCCGCGGCCATGCTGACAGG
TGTGGGCAAGGAGTATGCGCTTGGAGTGCAGATGGCTGTTGATGCAATTAAGCTGGTGAGTCGGGGAAAGCC
TTTGAGCGACGTTGCTTGTCCTCGATCTAGCGATACTTCCCGCCATTGTTATTTGTCCGCGGAGGAGTTGTGGA
AATATTGTCATGAGAGAAGAATTGAACTGGTGGATGATAAGAGGGTTGTGGAAATGCTGGTGGAGTCATTTG
CCCCGCTGGAAAAGAGGGAGGTGCTGCGAAAGGAGCTGGACGAGGCGGTGCAGGAATGGTATGTGAAGCA
AGAGGTGTGA

NNNNN = exon
NNNNN = intron